A Systematic Investigation of the Refinement Hypothesis:

Author: Karina Ashley Hamamouche

Boston College Electronic Thesis or Dissertation, 2019

A Systematic Investigation of the Refinement Hypothesis

Karina Ashley Hamamouche

A dissertation

submitted to the Faculty of

the department of Psychology

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

March, 2019

A SYSTEMATIC INVESTIGATION OF THE REFINEMENT HYPOTHESIS

Karina Ashley Hamamouche

Advisor: Sara Cordes, PhD

Throughout early childhood, children learn various symbolic systems to represent abstract concepts such as quantity. Yet it is unclear how the acquisition of symbols for quantity (e.g., number words; measurement concepts of "seconds", "minutes"… for time, etc.) may shape nonsymbolic representations of these quantities. While previous work hints at the possibility that acquiring numerical symbols refines numerical acuity (i.e., "refinement hypothesis"), these data are correlational in nature, making it impossible to assess causality. As such, experimental manipulations training the symbolic system are necessary in order to determine whether a causal relation exists. Moreover, these investigations have been limited to the domain of number, making it unclear if similar relations exist in continuous quantities, such as time and space. My dissertation tests whether the relation between symbolic and nonsymbolic abilities holds for the inherently continuous quantity of time, while also providing one of the first investigations of the refinement hypothesis outside of the domain of number. Results reveal that nonsymbolic and symbolic timing are related in childhood, both before and during formal instruction on temporal units of measurement (Experiment 1 & 2), but not in adulthood (Experiment 3). Further, I find no support for the refinement hypothesis: learning temporal symbols did not result in improved temporal acuity (Experiment 2), nor did shifting adults' symbolic mapping of time shape temporal acuity (Experiment 3). Similarly, learning labels for surface area did not enhance adults' spatial acuity (Experiment 4). Broader educational implications and areas of future investigation are also discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Words cannot express how grateful I am for the overwhelming amount of support I have received while completing my graduate work. None of this would have been possible without the following individuals:

To my advisor, Sara Cordes, for all of the time spent answering my questions, reading my manuscripts, and facilitating the development of my research program. Thank you for molding me into the researcher I am today and supporting my love of teaching. I cannot thank you enough.

To my undergraduate advisor, Neil Bohannon, for helping me discover my love of psychology and introducing me to psychological research.

To my committee, Elizabeth Kensinger, Marina Vasilyeva, and Joonkoo Park, for your unwavering support and insightful suggestions these past five years. Your guidance has made me a better scientist.

To Ehri Ryu, for the numerous statistical consulting sessions you provided me during my graduate career and especially while writing my dissertation.

To the Infant and Child Cognition Lab, in particular Tasha, Michelle, Sophie, Lindsey, Nadia, Stacee, Lindsey, Raychel, and Lauren, for creating a positive work environment, and influencing my work through both formal and informal conversations. A special thank you to my lab twin Sophie, for both programming and moral support throughout these past five years. I am so grateful to have had you by my side while experiencing the ups and downs of graduate school. I am also so thankful for the dedication and hard work of my research assistants without whom this work would have been impossible.

To the faculty, staff, and students at Pittsboro Primary School and the Conservatory Charter Lab, for welcoming me into your classrooms to conduct my dissertation work.

To my mom, for always finding a way to provide me with opportunities to grow and challenge myself. Thank you for never doubting my ability to finish my PhD and teaching me how to persevere when times get tough.

To Katia, for helping me interpret my findings through an educator's point of view and ensuring I never take life too seriously. Thank you for allowing me to work with your school and become Miss Hamamouche #2.

To my grandparents, your encouragement has been and will continue to be a driving force in my life.

To Eric, for believing in me when I did not believe in myself. Thank you for being my emotional support team and a sounding board for new research ideas, but most of all for always being my biggest fan.

CHAPTER 1: GENERAL INTRODUCTION

Symbols are culturally acquired representations that do not inherently convey information about the item they are representing. These symbols take many forms - written and spoken language, Arabic numerals, and even road signs. Given the importance of symbols in our everyday lives, it is not surprising that over the course of development, we learn various symbolic systems and regularly use them to convey abstract concepts, such as quantity (e.g., number is represented with Arabic numerals, time with seconds, minutes, hours…etc.). While research suggests that symbolic knowledge, such as language, shapes our perceptions of the world around us (e.g., Athanasopoulous, Wiggett, Derning, Kuipers, Thierry, 2009; Boroditsky, 2001; Boroditsky, 2009; Boroditsky, 2011; Boroditsky, Fuhrman, & McCormick, 2011; Casasanto, 2008; Casasanto, et al., 2004; Fausey, Long, & Boroditsky, 2009; Winawer, Witthoft, Frank, Wu, Wade, & Boroditsky, 2007), it is unknown whether the acquisition of quantitative symbols may shape approximate, nonsymbolic representations of these concepts. For example, does acquiring formal symbols for time (e.g., knowing exactly how long a second is) shape our ability to track temporal intervals?

Past work, particularly in numerical cognition, has identified nonsymbolic representations of number as the foundation upon which formal, symbolic math is built ("scaffolding hypothesis"; Barth, La Mont, Lipton, & Spelke, 2005; Gilmore, McCarthy, & Spelke, 2007; Holloway & Ansari, 2009; Mazzocco, Feigenson, & Halberda, 2011). That is, nonsymbolic representations of number mold our symbolic understanding of number. A second less explored possibility is that acquiring symbolic knowledge shapes approximate, nonsymbolic (i.e., not linguistically-dependent) representations of quantity ("refinement hypothesis"; e.g., Lyons, Bugden, Zheng, De Jesus, & Ansari, 2018; Matejko & Ansari, 2014; Mussolin, Nys, Content, & Leybaert, 2014; Shusterman, Slusser, Halberda, Odic, 2016; Suárez-Pellicioni & Booth, 2018). In my dissertation, I systematically investigate the refinement hypothesis in the domains of time and space.

Nonsymbolic Quantity Representation

Throughout the lifespan, humans and non-human animals are capable of representing nonsymbolic numerosities, durations, and spatial extents (e.g., surface area, cumulative area, etc.; Brannon & Terrace, 1998; Feigenson, Dehaene, & Spelke, 2004; Izard, Sann, Spelke, & Streri, 2009; Pica, Lemer, Izard, & Dehaene, 2004; Provasi, Rattat, & Droit-Volet, 2011; Brannon, Luz, & Cordes, 2006). These abilities are important for everyday activities such as learning language (Conway, Bauernschmidt, Huang, & Pisoni, 2010; Romberg & Saffran, 2010; Saffran, Aslin, & Newport, 1996), and decision-making (Lusardi, 2012; Peters, Västfjäll, Slovic, Mertz, Mazzocco, & Dickert, 2006; Reyna, Nelson, Han, & Dieckmann, 2009). Moreover, nonsymbolic temporal, numerical, and spatial acuity predict formal, symbolic mathematics (time, space, and number: Skagerlund & Träff, 2016; time and number: Odic, Lisboa, Eisinger, Olivera, Maiche, & Halberda, 2015; number and space: Lourenco, Bonny, Fernandez, & Rao, 2012; space: Bonny & Lourenco, 2015; number: Holloway & Ansari, 2009; Halberda, Mazzocco, & Feigenson, 2008; Mazzocco, et al., 2011; Starr, Libertus, & Brannon, 2013) and nonsymbolic temporal judgments are related to formal reading abilities (Casini, Pech-Georgel, & Ziegler, 2017; Cohen-Mimran & Sapir, 2007; Hood & Canlon, 2004; Plourde, Gamache, Laflamme, & Grondin, 2017; Tallal, 1980; for a review: Farmer & Klein, 1995).

Discrimination tasks, in which participants are asked to judge which of two numerical dot arrays (e.g., 4 versus 8 dots), durations (e.g., which of two ovals is displayed longer), or amounts (e.g., which of two arrays takes up more area or which of two lines is longer) is larger/longer, are commonly used to assess nonsymbolic representations of quantity. Numerous studies have revealed that over development and in non-human species, quantity discriminations adhere to Weber's Law, such that the speed and accuracy with which discriminations are made is dependent upon the ratio between the two quantities, rather than the absolute difference between them (Meck & Church, 1983; Moyer & Landauer, 1967; Stevens, 1957; for reviews: Gallistel, 1990; Feigenson, 2007). According to Weber's Law, discriminating 100 ms from 200 ms should be just as easy as discriminating 500 ms from 1000 ms because both comparisons involve a 1:2 ratio.

In addition to adhering to Weber's Law, representations of nonsymbolic quantities become more precise with age (e.g., Odic, Libertus, Halberda, Feigenson, 2013; Droit-Volet, Clément, Fayol, 2008). For instance, 6-month olds consistently detect a 1:2 ratio change in numerosity, duration, and surface area, but fail to detect a 2:3 ratio change (Brannon, et al., 2006; Brannon, Suanda, & Libertus, 2007; Xu & Spelke, 2000; Feigenson, 2007). Discrimination precision continues improving throughout development, such that the ratio that can be discriminated becomes narrower with age regardless of the quantity (see Odic, 2018). Although it is clear that quantity discriminations become more precise throughout development, the source of developmental change has been left to speculation. Maturation of neural mechanisms subserving these quantity representations (e.g., Lipton & Spelke, 2003) and/or burgeoning linguistic capacities, specifically pertaining to the domain of relevance (i.e., learning number words may shape approximate numerosity;

Shusterman et al., 2016; Posid & Cordes, 2015) have both been suggested as potential sources of increased precision. However, experimental tests of these possibilities have been lacking. As such, I investigate the possibility the acquisition of symbolic knowledge results in increased precision in nonsymbolic quantity judgments over the course of development (e.g., Lyons et al., 2018; Matejko & Ansari, 2014; Mussolin et al., 2014; Piazza, Pica, Izard, Spelke, & Dehaene, 2013; Shusterman et al., 2016; Suárez-Pellicioni & Booth, 2018).

Relation between Symbolic and Nonsymbolic Representations of Number

Scaffolding Hypothesis

Being able to track approximate quantities is critical for the development of symbolic thought. As such, many studies suggest that symbolic knowledge of number is acquired by mapping abstract symbols (e.g., Arabic numerals) onto the approximate number system - the analog magnitude system specifically used for representing number (Holloway & Ansari, 2009; Izard & Dehaene, 2008; Lipton & Spelke, 2005; Mazzocco, et al., 2011; Mundy & Gilmore, 2009). Evidence in favor of the scaffolding hypothesis comes from work exploring symbolic number comparisons. Similar to the nonsymbolic comparisons described above, symbolic number comparisons require participants to make judgments as to which of two simultaneously presented Arabic numerals is larger. Results of these studies have shown that the speed with which the larger of two Arabic numerals is selected is also ratio-dependent (Moyer & Landauer, 1967). In addition to behavioral evidence, neural and clinical data also reveal relations between symbolic and nonsymbolic representations of number. For instance, the intraparietal sulcus (IPS) is activated during both symbolic (e.g. Pinel, Dehaene, Riviere, & LeBihan, 2001) and nonsymbolic (e.g., Ansari, Dhital, & Siong, 2006) numerical processing. These direct parallels between symbolic and nonsymbolic comparisons have led researchers to conclude that symbolic representations are mapped onto nonsymbolic representations, at least in the domain of number.

Mapping symbolic representations onto nonsymbolic ones not only shapes our understanding of Arabic numerals, but also forms the foundation for more complicated mathematics. A substantial literature reveals that nonsymbolic representations of number predict later mathematics success (e.g., Bonny & Lourenco, 2015; Halberda, Mazzocco, & Feigenson, 2008; Starr, Libertus, & Brannon, 2013; for a review, see: Feigenson, Libertus, & Halberda, 2013; for meta-analyses: Chen & Li, 2014; Schneider, Beeres, Coban, Merz, Schmimdt, Stricker, & De Smedt, 2017). For example, numerical discrimination in early infancy is predictive of performance on standardized math assessments several years later (Starr et al., 2013), and preschoolers' approximate sense of number predicts their math abilities at age six (Halberda et al., 2008). These studies have been taken as further evidence that preverbal representations of number may form the foundation for later formal mathematics knowledge. This relation has also been found in adulthood, although less consistently (Libertus, Odic, & Halberda, 2012, but see Gilmore et al., 2013; Inglis, Attridge, Batchelor, & Gilmore, 2011; Skagerlund & Träff, 2016).

Additional evidence for this relation comes from training studies, in which a brief nonsymbolic addition or numerical comparison training, but not a comparable line length or brightness training, enhanced the speed with which children completed a symbolic math task (Hyde, Khanum, & Spelke, 2014; Khanum, Hanif, Spelke, Berteletti, & Hyde, 2016). Similarly, three to five-year olds from low income families who complete a nonsymbolic addition and subtraction training show select improvements in math compared to their peers who completed a memory training (Park, Bermudez, Roberts, & Brannon, 2016). In particular, math gains following approximate arithmetic training may be particularly beneficial for children who begin with low initial scores (Szkudlarek & Brannon, 2018). Training adults' approximate number sense also results in enhanced math abilities (Park & Brannon, 2013), although it less clear if this is a direct effect of increasing the precision with which we perceive number, or a broader manipulation of one's working memory abilities (Park & Brannon, 2014).

Refinement Hypothesis

While many studies support a causal relation such that nonsymbolic numerical abilities predict and shape symbolic ones, very few studies have tested the refinement hypothesis, or the idea that acquiring symbolic knowledge of number also impacts one's nonsymbolic numerical abilities. A small, but growing literature revealing math abilities predict later numerical acuity supports this possibility (Lyons et al., 2018; Matejko & Ansari, 2014; Mussolin et al., 2014; Piazza, Pica, Izard, Spelke, & Dehaene, 2013; Suárez-Pellicioni & Booth, 2018). In a related study, preschoolers' nonsymbolic numerical acuity improved dramatically after acquiring the cardinal principle (i.e., understanding that the final number in the count list indicates the set size) even while controlling for age (Shusterman et al., 2016), suggesting that the acquisition of numeric symbols resulted in improved nonsymbolic numerical acuity. Furthermore, children with dyscalculia, a developmental disorder that effects symbolic number processing, show nonsymbolic processing deficits (e.g., Mazzocco et al., 2011). This finding hints at the possibility that children who lack symbolic knowledge may experience less refinement of their nonsymbolic abilities.

While correlational research supports the refinement hypothesis, only a single study has investigated this relation experimentally. This work revealed that six sessions of symbolic arithmetic training did not impact adults' subsequent numerical acuity (Lindskog, Winman, & Poom, 2016). Given that the research investigating the refinement hypothesis has been largely correlational in nature, and several recent cross-lagged design studies have provided support for this hypothesis (e.g., Lyons et al., 2018; Mussolin et al., 2014), experimental manipulations are necessary to determine whether acquiring symbolic knowledge *causes* improvements in nonsymbolic acuity. My dissertation aims to shed light on the causal nature of this relation across development by experimentally manipulating individuals' symbolic knowledge of time (Experiments 2 & 3) and space (Experiment 4).

Why Symbols Shape our Perception: Evidence from Language

If there is support for the refinement hypothesis, one might question the underlying mechanism(s). That is, *why* might symbol acquisition impact nonsymbolic abilities? Borrowing from language literature, I propose that acquiring symbolic knowledge may (1) guide attention towards relevant cues and/or (2) facilitate memory encoding of relevant information (Boroditsky, 2001).

Heightened Attention

A plethora of work suggests that hearing a verbal label directs one's attention (e.g., Balaban & Waxman, 1997; Boutonnet & Lupyan, 2015; Lupyan & Ward, 2013). For instance, infants pay more attention to labeled compared to unlabeled objects (Balaban & Waxman, 1997), suggesting that hearing a word directs their attention towards the object. Similarly, in adulthood hearing a label enhances visual processing of that item (Boutonnet & Lupyan, 2015) by directing one's attention towards a previously unseen object (Lupyan

& Ward, 2013). For example, adults were quicker to find the letter e.g., "b" among a series of e.g., "p's" if they heard "bee" beforehand (Lupyan, 2008, see also Lupyan & Spivey, 2010). Given this work, it is possible that learning symbols to represent quantities may similarly heighten attention towards that quantity dimension. That is, one's attention towards e.g., time may be heightened after learning temporal units of measurement because temporal durations become more salient and relevant to other learning.

Memory Encoding

Likewise, a broad literature reveals memory encoding of information to be dependent upon one's language (e.g., Fausey & Boroditsky, 2011; Tosun, Vaid, & Geraci, 2013). For instance, there is a memory advantage for bilinguals if they encode and retrieve information in the same language (for a review: Schroeder & Marian, 2014). Other work shows that although English and Spanish speakers remember agents involved in intentionally events similarly, English speakers remember agents in accidental events better than Spanish speakers. It has been hypothesized that distinctions in the way each language encodes intentional versus accidental events subsequently impacts their memory of the event itself (Fausey & Boroditsky, 2011). Relatedly, Turkish, but not English, speakers remember information from first hand sources better than non-firsthand sources, likely because the Turkish language requires speakers to specify the knowledge source, whereas English does not (e.g., Tosun et al., 2013). Thus, learning quantity symbols may make it easier to encode information about that quantity in memory, resulting in enhanced processing of that quantity and thus refinement of nonsymbolic representation.

While investigations with language provide some insight into how symbolic knowledge may shape our perception of the world around us, it is unknown whether the same mechanism(s) may explain *how* the acquisition of quantitative symbols shapes nonsymbolic representations of quantity. Although Shusterman and colleagues (2016) argue that acquiring the cardinal principle (i.e., symbolic knowledge of number) likely heightened children's attention to number, resulting in improved numerical acuity, no work has experimentally investigated this possibility. Thus, the final experiment of my dissertation explores whether the acquisition of spatial symbols leads to heightened attention towards spatial dimensions and/or enhanced memory encoding of spatial dimensions.

The Current Dissertation

Across four experiments, I investigate the broad question of how the acquisition of formal, symbolic representations of quantity shapes our nonsymbolic quantity representations. Although the domain of number has been inundated with investigations exploring the relation between symbolic and nonsymbolic representations (e.g., Halberda, et al., 2008; Starr, et al., 2013), it is unknown whether these relations hold in other quantities such as time and space. Studying these relations in time and space is particularly fruitful for several reasons. First, both nonsymbolic number and numeric symbols are discrete in nature, leaving it is unclear what role these structural similarities may play in promoting a relation between symbolic and nonsymbolic quantities. If structural isomorphisms support this relation, then it might be more difficult for discrete symbols to impact nonsymbolic representations of continuous quantities, such as time or space. Despite this, no work has explored whether a similar relation holds when mapping discrete symbols onto inherently continuous quantities.

Secondly, children acquire count words early in development (i.e., preschool years), a period during which children experience several gains in relevant cognitive skills (e.g., executive functions; Epsy, 1997) and neurological maturity (Emerson & Cantlon, 2015), thus making it unclear how much these other developmental factors may contribute to changes in nonsymbolic acuity improvements. By exploring the acquisition of temporal and spatial symbols, which take place at a later point in development, I aim to disentangle the effects of maturity and other developmental milestones on increasing quantity acuity throughout development.

In Experiments 1 and 2, I first investigate whether a relation exists between children's understanding of temporal units of measurement (i.e., symbolic understanding of time) and their nonsymbolic temporal acuity. After the relation has been established, I then test the possibility that acquiring symbols for time refines children's temporal acuity (Experiment 2). Extending this work to adults, in Experiment 3, I determine whether a similar correlation between nonsymbolic and symbolic temporal abilities exists beyond childhood. Because adults already have a substantial understanding of symbolic time, I cannot test the causal nature of this relation by teaching adults about temporal symbols. Instead, I shift their symbolic mapping of time through feedback to determine whether this impacts their subsequent temporal acuity. Finally, in Experiment 4, I study the relation in the domain of space (e.g., surface area), to (1) provide converging evidence in a distinct quantitative domain and (2) assess whether linguistic symbols may serve to heighten salience and/or provide more efficient encoding of the relevant quantity. Together, my dissertation serves as one of the first systematic experimental investigation of the acquisition of quantity symbols.

CHAPTER 2: KNOWLEDGE OF FORMAL TIMING SYMBOLS IS RELATED TO INDIVIDUAL DIFFERENCES IN TEMPORAL PRECISION

Abstract

Throughout the lifespan, we are capable of representing quantities in the absence of language, or nonsymbolically. Additionally, over the course of development, we learn many symbolic measurement systems for representing quantities such as time and number. Despite substantial evidence of a relation between the acquisition of symbolic and nonsymbolic numerical acuity (see Halberda et al., 2008), no work has explored whether a similar relation exists between understanding temporal units of measurement and timing precision. That is, does a child's understanding of words like "second", "minute", and "hour" have any relation to their ability to tell which of two events lasted longer? Six and seven-year-old children (N = 102, M_{age} = 83.44 months, 52 females), who are in the process of learning temporal units of measurement, completed a temporal discrimination task (assessing nonsymbolic temporal acuity) and a symbolic timing assessment. Results revealed a positive correlation between children's nonsymbolic temporal acuity and their understanding of temporal units of measurement. Importantly, this correlation held when controlling for age and numerical acuity, suggesting a unique relation between children's temporal acuity and their understanding of temporal units of measurement. This study is the first to show a relation between symbolic and nonsymbolic representations of time.

Over development, children learn various measurement systems (i.e., symbol systems) to represent abstract quantity concepts (i.e., number, time, space, etc.). While these symbols (e.g., Arabic numerals for number, units of measurement for time, etc.) allow for the exact measurement of quantity, it is unknown whether an understanding of these systems may be related to our ability to perceive and/or discriminate these quantities in the absence of symbols (often referred to as nonsymbolic representations of quantity). Despite a plethora of research demonstrating a link between nonsymbolic numerical acuity and symbolic representations of number (i.e., counting, mathematics achievement; Bonny & Lourenco, 2015; Chen & Li, 2014; Libertus, Feignson, & Halberda, 2011; Mazzocco, Feigenson, & Halberda, 2011a; Mazzocco, Feigenson, & Halberda, 2011b; Halberda, Feigenson, & Mazzocco, 2008; Schneider et al., 2016), it is unknown whether a similar relation holds for other quantities such as time. In the current study, we tested the relation between nonsymbolic and symbolic representations of time in 6-7-year olds, children who are in the process of learning temporal units of measurement in school.

The current study is important for several reasons. First, while age-related increases in temporal precision are well-documented (see Brannon, Suanda, & Libertus, 2007; Droit-Volet, 2013; Odic, 2018), the source of these changes is unknown. One possibility is that brain maturation and/or other general cognitive abilities may contribute to more precise quantity tracking (e.g., Lipton & Spelke, 2003). However, it has also been suggested that the acquisition of quantity symbols may play a role in altering the precision with which quantities are tracked nonsymbolically (e.g., Pica, Lemer, Izard, & Dehaene, 2004; Posid & Cordes, 2015; Shusterman, Slusser, Halberda, & Odic, 2016). While the current investigation is correlational and thus cannot speak to the question of causality (i.e., whether nonsymbolic timing abilities form the foundation for learning about symbolic representations of time, or alternatively, whether learning about temporal symbols may shape our nonsymbolic timing abilities), this study sets the stage by providing the first data to establish the existence of such a relation. Secondly, although a relation between symbolic and nonsymbolic *numerical* abilities has been established, it is important to note that both numerical symbols and nonsymbolic representations of number are discrete in nature, which may facilitate this relation. If structural isomorphisms underlie this link, then it may be more difficult to find a similar association between temporal acuity and an understanding of discretized temporal units of measurement (e.g., an understanding of seconds, minutes, etc.), as nonsymbolic representations of time are inherently continuous. Lastly, recent work identifying links between timing abilities and formal mathematics (Hamamouche & Cordes, 2019; Kramer, Bressan, & Grassi, 2011; Odic et al., 2016; Skagerlund & Träff, 2016) further emphasizes the importance of understanding children's developing temporal abilities. Thus, understanding if a relation exists between symbolic and nonsymbolic timing may provide hints at the source of age-related increases in timing precision, shed light on the mechanisms underlying the relation between numerical acuity and verbal counting, and also inform our understanding of temporal development in childhood.

Nonsymbolic Temporal Processing

Throughout the lifespan, humans and non-human animals are capable of representing time in the absence of language (e.g., Brannon et al., 2007; Droit-Volet, 2013; Droit-Volet & Wearden, 2001; Meck & Church, 1983; Platt & Davis, 1983; Provasi, Rattat, & Droit-Volet, 2011; van Marle & Wynn, 2006). This ability, referred to as our *nonsymbolic* representations of time, is considered the most basic and intuitive form of quantity representation. Critically, nonsymbolic representations of time do not rely on linguistic or symbolic information. These representations are often measured using discrimination tasks, in which participants are asked to determine which of two durations was longer.

Substantial work has characterized our nonsymbolic timing abilities. Like other quantities, nonsymbolic representations of time adhere to Weber's Law (Gibbon, 1977; Hamamouche & Cordes, 2019); such that the speed and accuracy with which discriminations are made is dependent upon the ratio between the two durations, rather than their absolute difference (Stevens, 1957). For example, the ease with which 1 second is discriminated from 2 seconds is comparable to that with which 5 seconds is discriminated from 10 seconds because both comparisons involve a 2-fold ratio change. Moreover, our ability to make these discriminations becomes more precise with age (e.g., Odic, 2018; Odic, Libertus, Halberda, Feigenson, 2013; Droit-Volet, Clément, Fayol, 2008). Six-month old infants reliably detect a 2-fold change in duration of two stimuli, but fail to detect a 1.5-fold change. Within a few months, however, their timing abilities improve such that 10-month old infants can detect a 1.5-fold change (Brannon et al., 2007; Feigenson, 2007; vanMarle & Wynn, 2006). The precision with which we discriminate time continues to improve throughout childhood, such that the ratio that can be discriminated approaches one with age (see Odic, 2018).

Although it is clear that quantity discriminations become more precise throughout development, the source of developmental change is unclear. Maturation of neural mechanisms subserving these quantity representations (e.g., Lipton & Spelke, 2003) and/or burgeoning linguistic capacities, specifically pertaining to the domain of relevance (e.g., learning temporal units of measurement; Shusterman et al., 2016; Posid & Cordes, 2015) have both been suggested as potential sources of increased precision, at least in the domain of number. In this study, we explore the possibility that the acquisition of temporal units of measurement may be related to increased precision in temporal acuity over the course of development (see Piazza, Pica, Izard, Spelke, & Dehaene, 2013).

Symbolic Representations of Time

Although nonsymbolic representations of time have been extensively explored, few studies have assessed children's symbolic understanding of time (i.e., understanding of temporal units of measurement). Just as temporal acuity increases over development, an understanding of temporal units of measurement also increases, but through learning. For example, children's use of temporal vocabulary (e.g., words such as "seconds", "minutes", etc.) increases particularly during the early elementary years (e.g., Ames, 1946; Bradley, 1947; Harrison, 1934; Oakden & Sturt, 1922; Tillman & Barner, 2015). However, despite using temporal vocabulary early on, children often lack a full understanding of the words' meanings (Ames, 1946; Shatz, Tare, Nguyen, & Young, 2010). For example, children as young as four years old are able to successfully order durations associated with these words (e.g., a minute is less than an hour) without knowing the approximate duration associated with each temporal word (Tillman & Barner, 2015; see also Friedman, 1977). In addition to understanding temporal vocabulary, children also learn to use clocks and stopwatches to track time during early childhood. Much like understanding the meaning of temporal words, this process takes several years to master (Friedman & Laycock, 1989).

It is unknown whether a relation exists between temporal acuity and an understanding of temporal units of measurement; however, one study hints at the likelihood of this relation. Arlin (1990) asked children whether clocks move faster, slower, or do not change speed while they sleep – an indicator of children's symbolic understanding of time. Then, children watched as two puppets danced for the same amount of time (15 seconds), but at different speeds (140 beats per minute versus 40 beats per minute) and were asked whether one puppet danced longer than the other or if the puppets danced for the same amount of time (a proxy for nonsymbolic timing). Results revealed that children who understood that the clock maintains its speed at night were more likely to say that the puppets danced for the same amount of time, suggesting that a relation may exist between symbolic and nonsymbolic timing (Arlin, 1990). While these data point to a possible relation between representations of time, the tasks used were not pure measures of nonsymbolic and symbolic timing. Thus, the current study is the first to characterize this relation by using a standard nonsymbolic temporal discrimination task and a broader assessment of children's symbolic timing abilities.

Relations between Symbolic and Non-Symbolic Quantity Representations

While the relation between symbolic and nonsymbolic *timing* has been left untested, numerous studies have reported a link between symbolic and nonsymbolic *numerical* acuity (e.g., Chen & Li, 2014; Schneider et al., 2016). Numerical discriminations in early infancy have been found to predict performance on standardized math assessments several years later (Starr et al., 2013), and preschoolers' approximate sense of number predicts their math abilities at age six (Halberda et al., 2008). This relation holds even in adulthood (e.g., Libertus, Odic, & Halberda, 2012), although less consistently so (Gilmore, McCarthy, & Spelke, 2013; Inglis, Attridge, Batchelor, & Gilmore, 2011; Skagerlund & Träff, 2016).

Importantly, while a well-documented relation exists between symbolic and nonsymbolic numerical abilities (e.g., Bonny & Lourenco, 2015; Halberda, et al., 2008; Starr, et al., 2013; for a review, see: Feigenson, Libertus, & Halberda, 2013), it is unclear whether this relation is specific to number, or if it holds for other quantities. Thus, in the present study, we investigate whether a relation exists between nonsymbolic and symbolic timing. That is, does the acquisition of formal measurement terms in the timing domain (such as learning how long a "second" or a "minute" lasts; hereafter referred to as "temporal units of measurement") relate to our ability to perceive the duration of a stimulus? Unique to number, both symbolic and nonsymbolic representations of number are discrete in nature, such that structural similarities may support the link between the two. Such structural isomorphisms fail to hold for non-discrete quantities, such as time, making the existence of a relation between the two less straightforward. Despite this, recent work suggests that this relation holds in the domain of space (Lauer & Lourenco, 2016; Lourenco & Bonny, 2017). For example, Lourenco and Bonny (2017) found children's ability to discriminate cumulative surface area to be related to a formal understanding of geometry. Importantly, nonsymbolic representations of space are continuous, whereas spatial units of measurement are discrete. Moreover, there is evidence that children are not only able to map between continuous and discrete magnitudes (see Babai, Nattiv, & Stavy, 2016), but they are also capable of employing strategies used to compare continuous magnitudes to situations in which magnitudes are discretized (Boyer & Levine, 2015). Given these findings, we predict a relation between nonsymbolic and symbolic timing.

Additionally, many similarities between temporal, spatial, and numerical processing (see Walsh, 2003) also lend support for our hypothesis that children's nonsymbolic and symbolic timing abilities are correlated. For instance, numerical, spatial, and temporal discriminations adhere to Weber's Law (e.g., Brannon, Lutz, & Cordes, 2006; Izard, Sann, Spelke, & Streri, 2009; Provasi et al., 2011) and precision in these domains increases over development (Odic, 2018). Moreover, the precision of temporal, numerical, and spatial discriminations follow comparable developmental trajectories in infancy (for a review: Feigenson, 2007). In addition, temporal, numerical, and spatial acuity predict math achievement (see Lourenco, Bonny, Fernandez, & Rao, 2012; Odic et al., 2016). These similarities and others have led many to contend that a single common magnitude system is responsible for processing quantities (Cantlon, Platt, & Brannon, 2009; Walsh, 2003). That is, it has been posited that it is no coincidence that temporal, numerical and spatial processing show similar behavioral signatures because all three quantitative magnitudes are thought to be represented via a common neural locus and neural code. Thus, those in favor of the common magnitude system would also predict a relation between nonsymbolic and symbolic timing given (1) the similarities between numerical, spatial, and temporal processing and (2) the presence of relations between nonsymbolic and symbolic in other quantitative domains.

Even in the absence of a common magnitude system wherein each magnitude relies upon a unique system (see Hamamouche & Cordes, 2018; Odic, 2018), one might still find reason to predict a relation between nonsymbolic and symbolic timing. For one, individuals who have a greater symbolic understanding of time may be better able to encode time. Moreover, people with a greater understanding of temporal units of measurement may employ more efficient strategies when timing, again leading to enhanced nonsymbolic timing abilities. Thus, regardless of the system(s) responsible for quantity processing, we predict a relation between nonsymbolic and symbolic timing in children.

While we predict a relation between nonsymbolic and symbolic timing, what if the relation does not exist? Such a finding would be surprising, as it would suggest that our acquisition of symbolic time is entirely unrelated to our timing abilities. While the null results may be due to noise in the data, this lack of relation could also point to a striking dissociation between symbolic and nonsymbolic timing abilities, while possibly leading some to question current theories of numerical development. Moreover, a lack of relation would also hint at differences in temporal and numerical processing.

The Current Study

In the current study, we investigated whether a relation exists between nonsymbolic and symbolic timing abilities in 6-7-year old children. Because people tend to master temporal symbols by adulthood, we purposely explored this relation in a population likely to show maximal variability in temporal symbol understanding – children who are in the process of learning temporal symbols in school. Children completed a nonsymbolic timing task (Temporal Discrimination), and two assessments of symbolic timing (Temporal Units of Measurement Questionnaire and a Temporal Estimation task).

Methods

Participants

One hundred and two 6-7-year olds (M_{age} = 83.44 months, 52 females) participated. An additional seven children completed only a single task in its entirety and were not included in the analyses. This age range was chosen to focus on a period of development during which children begin acquiring formal units of measurement for time in the classroom (National Governors Association Center for Best Practices, 2010). In particular, according to the Common Core standards, first graders are learning how to tell time to the nearest hour and half hour. By second grade, students should be able to recognize the relations between different temporal units of measurement (e.g., a minute is longer than a second). As such, this age range was chosen so as to maximize variability in symbolic understanding across our sample.

Children from the greater Boston area were recruited at afterschool programs (n = 7), local museums or parks (n = 54), or during a short visit to the lab (n = 41). The battery of tasks took approximately 15 minutes to complete. Required sample size was based on a priori calculations in G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) requiring 0.80 power for detecting a medium effect size on central analyses. All methods were approved by the Boston College IRB.

Procedure

After parental consent was obtained, the experimenter gave parents a Parent Timing Questionnaire (modeled after LeFevre, Skwarchuk, Smith-Chant, Fast, & Kamawar, 2009). Like LeFevre et al. (2009), we included questions about the frequency of using calendars and dates (separated into two questions as opposed to a single question in LeFevre et al., 2009), and wearing a watch. We also added new questions that related to reading digital and analog clocks, discussing time, and using a timer[1]. Because responses

[1] In our sample, we found higher rates of using calendars and dates (M = 3.79, SD = 1.23 compared to LeFerve et al., 2009, M = 3.0, SD = 1.1, $t(244) = 5.27$, $p < .001$), and wearing watches (M = 1.7, SD = 1.18 versus LeFevre et al., 2009 M = 1.2, SD = 1.4, $t(244) = 2.93$, $p < .01$), likely due to differences in the age range our samples.

to this questionnaire were not correlated with any of the child timing measures, we concluded that it was not an accurate depiction of the children's timing abilities and thus excluded it from the main analyses.

Children were tested in a quiet testing space. Children completed the following tasks in a single session in a set order: (1) temporal discrimination (assessment of nonsymbolic timing abilities), (2) temporal units of measurement Questionnaire (assessment of symbolic timing abilities), (3) temporal estimation task (assessment of symbolic mapping), and (4) numerical discrimination Task (assessment of nonsymbolic numerical abilities).

(1) Temporal discrimination (modeled after Hamamouche & Cordes, 2019)

The temporal discrimination task served as an assessment of the children's nonsymbolic temporal acuity. Children saw a cartoon chicken and cow presented side-by-side on the computer screen. Children were told that the chicken and cow were going to take turns playing musical instruments for a certain amount of time and their task was to decide which animal played their instrument longer. On every trial, the chicken (always located on the left) played its trumpet first, followed by a 1500 ms inter-stimulus interval, and then the cow (always located on the right) played its trumpet. When each animal played their trumpet, children heard a sound accompanied by the presence of a box around the relevant animal (both a visual and auditory cue). The trumpet of each animal made a unique sound, such that the chicken's trumpet always sounded like a horn and the cow's trumpet sounded like a train. Children indicated which animal played the trumpet longer by pointing and the experimenter recorded the response by pressing specified keys on a keyboard.

Children completed eight practice trials, during which children were encouraged to ask the experimenter questions. The practice trials were identical to the test trials except that the ratio between the two durations presented was 3.2. In order to ensure that children could clearly hear the stimuli, after completing the practice trials, the experimenter asked all children to put headphones on for the remainder of the task. Although the experimenter could not hear the sound of the discrimination stimuli, (s)he could still see the box appearing around each character, and thus the experimenter was not naïve to the stimuli.

During test, the durations of the stimuli ranged from 200 ms – 4000 ms. The pairs of durations presented on every trial differed by one of four ratios: 1.25, 1.5, 2, 2.25 (modeled after Hamamouche & Cordes, 2019; Odic et al., 2016). The cow played the instrument longer on exactly half of all trials. In test, children completed 32 trials: 4 ratios (1.25, 1.5, 2, 2.25) x 8 times each, with all manipulations intermixed in the same block of trials[2]. All stimuli were presented and data were recorded using a Xojo program on a Mac laptop.

(2) *Temporal units of measurement questionnaire (modeled after Oakden & Sturt, 1922 and Tillman & Barner, 2015)*

This was a measure of children's temporal symbol understanding. An experimenter read children 13 questions relating to how long it takes to do everyday activities (e.g., "How long does it take you to brush your teeth?"), reading analog and digital clocks, and identifying which duration is longer (e.g., "Which is longer a minute or a second?"). The

[2] A programming error led 19 participants to only receive 6 iterations of a 2-fold ratio change (and 2 iterations of a 1.75 change) on the time discrimination task. While these trials were included in analyses involving overall time discrimination performance, these trials were excluded in analyses concerning ratio effects.

experimenter read each question to the child and recorded their responses using pen and paper. The measure showed strong internal consistency, $\alpha = .861$. See Appendix A.

(3) Temporal estimation (modeled after Bueti, Walsh, Frith, & Rees, 2008)

This task served as a second assessment of children's symbolic mapping abilities. On a laptop using a Xojo program, children were asked to hold the left arrow key on a keyboard for a set amount of time (5, 9, 14 seconds; order counterbalanced). On each trial, the experimenter prompted the child by saying, "Hold down this key [pointing to the left arrow key] for [5, 9, or 14] seconds." The experimenter did not suggest strategies to the child; however, if the child began using a strategy such as counting, they were not stopped.

(4) Numerical discrimination: Panamath (Halberda et al., 2008)

The numerical discrimination task was added to the battery of tasks half-way through data collection, thus only a subset of children ($n = 59$) completed the task. We included this to assess whether discrimination performance in general, or temporal discrimination in particular, was related to an understanding of temporal units of measurement. During the numerical discrimination task, children saw two separate, but simultaneously presented, dot arrays containing either blue or yellow dots on the computer screen and were asked which color contained more dots. Children completed 88 trials in which dot arrays ranged from 5-21 dots. All dot arrays were displayed for 1951 ms and the two arrays differed by one of the following ratios: 1.28, 1.47, 1.75, 2.75. In half of the trials, the cumulative area of each dot array was equated, such that the larger dot array contained smaller individual dots on average. In the other half of the trials, the dots in both arrays were the same size on average. As such, the larger numerosity also contained a larger surface area. Children stated their answer out loud and the experimenter recorded their responses by pressing corresponding keys on the keyboard.

After completing all tasks, children and their parents were debriefed, and children picked out a prize for participation.

Data Coding and Analyses

Temporal units of measurement questionnaire

Accuracy (percent correct) was the dependent variable for the temporal units of measurement questionnaire. Responses indicating uncertainty (i.e., “I don’t know”, “I haven’t learned that yet”) were counted as incorrect. One child did not complete the temporal units of measurement questionnaire.

Temporal discrimination

Percent correct was the dependent variable for the temporal discrimination task. One child’s score was three standard deviations below the mean, and thus was excluded from the analyses. The data from one child was not saved due to a computer error, and one child showed a side bias by always choosing one of the characters, and thus, these children’s responses were not included in the analyses.

Temporal estimation

Absolute average error was calculated and used as the dependent variable for the temporal estimation task. Error was calculated by taking the absolute value of the difference between the child’s response to each duration and the actual duration to be estimated. This value was then divided by the duration to be estimated. Then, the average error across the three trials was calculated. Two participants did not understand the task and thus were excluded from the analyses. The average error from two participants was more than three standard deviations above the mean, and these participant's data was excluded from analyses involving temporal estimation. Four children did not complete this task (time constraints, wanting to discontinue participation, etc.).

Numerical discrimination

Percent correct was the dependent variable for numerical discrimination task. Two children had overall scores more than three standard deviations below the mean, and thus their performance was excluded. After implementing the task, 11 children (n = 7 due to time constraints, n = 4 program error) did not complete the numerical discrimination task, and thus their scores were not included. One additional child was excluded due to parental interference during the numerical discrimination task.

Results

See Table 2.1 for average performance on each task. Children performed above chance on the temporal discrimination ($t(98) = 23.84$, $p < .001$, $N = 99$) and the numerical discrimination ($t(44) = 32.78$, $p < .001$, $N = 45$). Table 2 shows the correlations between tasks.

Temporal discrimination	**Temporal units of measurement questionnaire**	**Temporal estimation error**	**Numerical discrimination**
79.48% (1.24) $N = 99$	76.00% (1.81) $N = 101$	.54 (.03) $N = 94$	91.74% (1.27) $N = 45$

Table 2.1. Mean performance (standard error in parenthesis) on each task.

	1.	2.	3.	4.
1. Temporal discrimination		.499** $N = 98$	-.260* $N = 92$	.321* $N = 42$
2. Temporal units of measurement questionnaire	.469** $N = 98$		-.319** $N = 97$	.410** $N = 45$
3. Temporal estimation	-.265* $N = 92$	-.376** $N = 94$		-.070 $N = 44$
4. Numerical discrimination	.285 $N = 42$	.450** $N = 45$	-.124 $N = 45$	

** $p < .01$, * $p < .05$

Table 2.2. Correlations between performance on tasks (below diagonal), controlling for Age (above diagonal).

Ratio Effects for Temporal and Numerical Discrimination

Previous research has reported ratio effects in temporal and numerical discriminations, such that larger ratios are easier to discriminate (indicative of Weber's Law; see Gibbon, 1977; van Marle & Wynn, 2006). To test for ratio effects in our temporal data, we conducted a repeated measures ANOVA on performance on trials for each of the four ratios in our temporal discrimination task. There was a significant effect of Ratio, $F(3, 294) = 44.47$, $p < .001$, $\eta_p^2 = .312$, $N = 99$, such that performance was worse on smaller ratios compared to larger ratios (*p's* < .001, except for the comparison between the two easiest ratios, $p = 1.0$; See Figure 2.1a). A comparable repeated measures ANOVA testing for ratio effects in our numerical discrimination data also revealed a main effect of Ratio, $F(3, 132)$, = 33.49, $p < .001$, $\eta_p^2 = .432$, N = 45, indicating better performance on trials consisting of a larger ratio (all *p's* < .01, except for the comparison between the two middle ratios, $p = .08$; See Figure 2.1b).

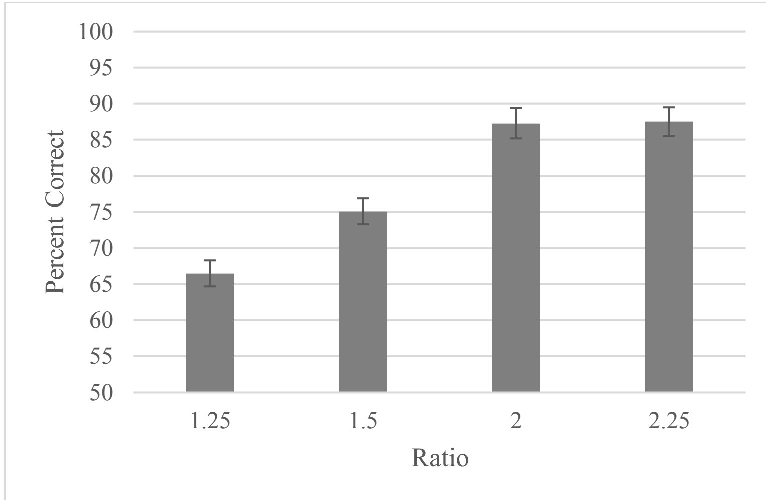

Figure 2.1a. Performance on the Temporal Discrimination task by Ratio (N = 99). Error bars represent standard error.

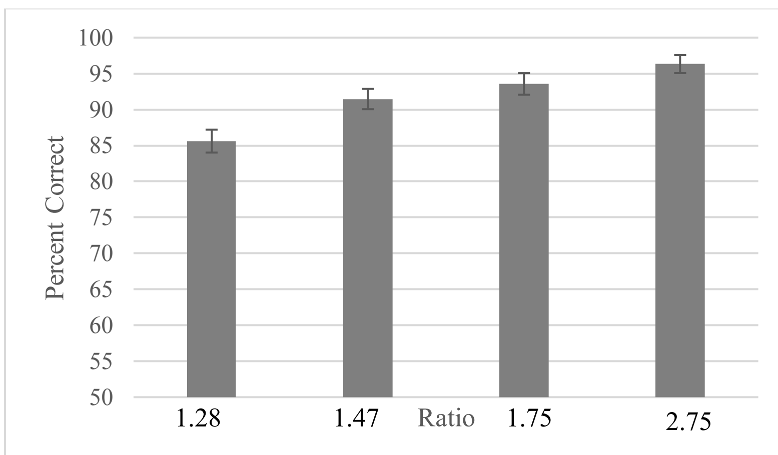

Figure 2.1b. Performance on the Numerical Discrimination task by Ratio (N = 45). Error bars represent standard error.

Temporal Estimation Performance

Children's estimates increased with increasing duration on the temporal estimation task ($F(2, 186) = 27.36, p < .001, \eta_p^2 = .227, N = 94$; all p's $< .01$). Despite this, the amount of error was comparable regardless of the duration children were asked to produce ($F(2, 186) = 1.94, p = .15, \eta_p^2 = .020$).

The Relation between NonSymbolic and Symbolic Timing in Childhood

Performance on our two measures of children's symbolic timing understanding, the temporal estimation task and the temporal units of measurement questionnaire, were correlated, $r = -.376$, $p <. 001$, $N = 94$. As such, we calculated a composite symbolic timing score by averaging children's performance on the temporal units of measurement questionnaire and the temporal estimation task to be used for all analyses. Because a higher score on the temporal units of measurement questionnaire was indicative of better performance, but a higher error score on the temporal estimation task was indicative of lower performance, we calculated a z-score for each task separately, and then multiplied the temporal estimation z-score by -1. Finally, we averaged the two z-scores into a single composite symbolic timing score. Five participants only had one of the symbolic timing measures, and thus were not included in the analyses involving the composite symbolic timing score. All analyses remain consistent when considering only one of the symbolic timing measures, except where indicated.

Previous work has shown that both nonsymbolic and symbolic abilities improve with age (Brannon et al., 2007; Tillman & Barner, 2015; van Marle & Wynn, 2006). To test this, we conducted separate correlations between age (in months) and performance on the temporal discrimination task, and age and performance on the numerical discrimination task. We also conducted a comparable correlation with age and our composite symbolic timing score. While age was correlated with symbolic timing ability, $r = .390$, $p < .001$, $N = 94$, it was not correlated with nonsymbolic timing abilities[3], $r = .045$, $p = .657$, $N = 99$, or nonsymbolic numerical abilities, $r = .277$, $p = .066$, $N = 45$.

[3] Based on other work suggesting a distinction in performance on sub- and supra-second timing emerges in childhood (Hamamouche & Cordes, 2019), we explored whether discrimination performance differed between sub and supra second ranges in our study; however, they did not, $p >.26$, $N = 100$.

Given prior work demonstrating a relation between numerical acuity and symbolic math (e.g., Halberda et al., 2008; Starr et al., 2013), it was predicted that children with a greater symbolic understanding of time would also have more precise nonsymbolic representations of time. Specific to this hypothesis, we conducted a correlation between children's symbolic and nonsymbolic timing abilities. The correlation was significant, $r = .450$, $p < .001$, $N = 92$, suggesting that greater temporal precision is related to a stronger understanding of temporal symbols. We next conducted a partial correlation controlling for age (in months) to test the relation between symbolic and nonsymbolic timing in childhood independent of age-related changes. Importantly, the correlation remained significant, $r = .464$, $p < .001$, $N = 92$, suggesting that this relation holds above and beyond age related changes[4]. See Figure 2.2.

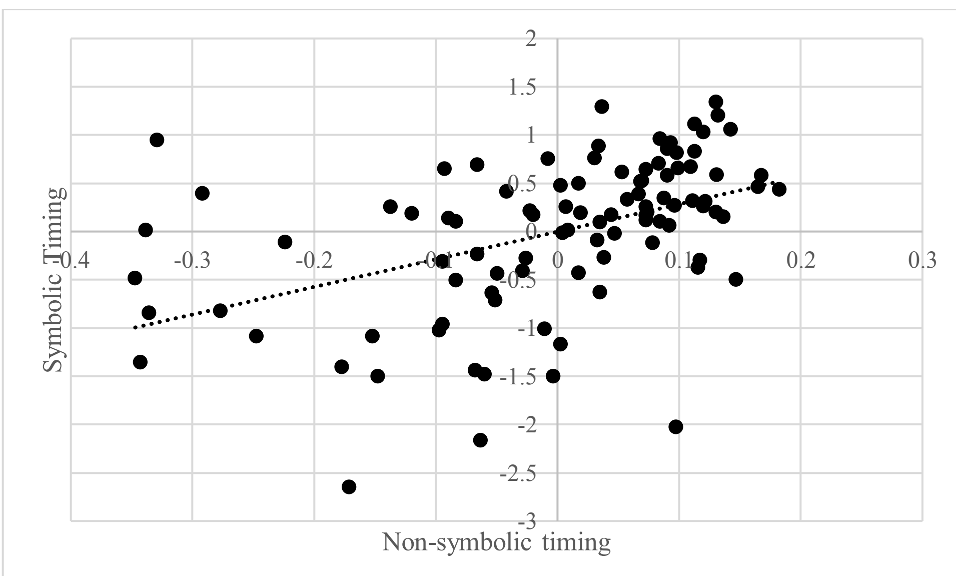

Figure 2.2. A strong positive relation exists between nonsymbolic timing abilities and symbolic timing abilities when controlling for age.

[4] This correlation remains significant when controlling for pairwise comparisons.

To assess whether symbolic timing abilities predict nonsymbolic timing precision, even when controlling for differences in developmental maturity (i.e., age) and basic discrimination capacities, we conducted a hierarchical regression. In the first step, we entered age and numerical discrimination as a predictor of nonsymbolic timing. Although the model did not reach significance, $F(2, 39) = 2.670$, $p = .082$, numerical discrimination (B = .006 $p = .041$), but not age (B = -.004, $p = .195$) predicted nonsymbolic timing. Adding symbolic timing into the second step added additional variance (R^2 change = .140, p = .011), allowing the model to reach significance, $F(3, 38) = 4.462$, $p = .009$. Symbolic timing (B = .060, $p = .011$) significantly predicted nonsymbolic timing, while age (B = -.005, p = .065) and numerical discrimination (B = .005, $p = .070$) marginally predicted nonsymbolic timing[5]. See Table 2.3. Thus, even when controlling for Age and for performance on a task with comparable demands in the domain of number, nonsymbolic timing was uniquely related to an understanding of temporal units of measurement.

Step 1:	B	SE B	β	p-value
Age	-.004	.003	-.202	.195
Numerical Discrimination	.006	.003	.324	.041
Step 2:				
Age	-.005	.003	-.275	.065
Numerical Discrimination	.005	.003	.267	.070
Symbolic Timing	.060	.022	.387	.011

Note: Only the second model was significant, $p = .009$. The R^2 change for Model 2 (.140, $p = .011$) reached significance.

Table 2.3. Hierarchical regression showing the unique relation between symbolic and nonsymbolic timing.

[5] Adding performance on the temporal estimation task alone did not contribute additional variance (R^2 change = .033, $p = .230$) to the model, and both models remined non-significant, *p's* > .05.

Discussion

While the field of mathematical cognition has been immersed with investigations of the relation between symbolic and nonsymbolic number and to a lesser degree, in space as well (number: Gilmore et al., 2010; Halberda et al., 2008; Libertus et al., 2011; Mazzocco et al., 2011; Starr et al., 2013; space: Lauer & Lourenco, 2016; Lourenco & Bonny, 2017), no work has explored this relation in the domain of time. Our study is the first to report a relation between symbolic and nonsymbolic timing abilities in children. As predicted, we found that nonsymbolic temporal acuity, as assessed by a temporal discrimination task, was significantly correlated with an understanding of temporal units of measurement in childhood. Importantly, these findings held even when controlling for age, revealing that these relations could not be accounted for by maturational changes that are expected to occur over the course of development. While investigations in numerical cognition have found this relation across a wide age span, we chose to test only 6-7-year olds as this is the age at which temporal units of measurement are taught in the classroom and thus the point at which the greatest changes in symbolic knowledge occur. Future work testing the consistency of this relation across a larger age range may be important for further characterizing this link. While one might expect this relation to exist in adults, a growing body of evidence has questioned the presence of a relation between adult's symbolic and nonsymbolic representations of number (e.g., Gilmore et al., 2013; Inglis et al., 2011; Skagerlund & Träff, 2016). Moreover, adults are likely to adapt timing strategies, such as counting, that may not rely upon nonsymbolic representations of time, making it less likely this relation might hold in adulthood.

Our analyses also revealed that the relation between symbolic and nonsymbolic timing abilities held strong in a subset of participants even when controlling for numerical discrimination abilities and age, indicating that the link between symbolic and nonsymbolic timing abilities is unlikely to be dependent upon basic maturation and/or more domain-general task demands (i.e., decision-making, perceptual acuity, and/or working memory demands). Although previous work has linked domain-general abilities like information processing speed and working memory to explicit timing performance (Droit-Volet & Zélanti, 2013a, b; Zélanti & Droit-Volet, 2011), our analyses suggest that these domain general abilities – likely implicated in performance in both the numerical and temporal discrimination tasks – are not the source of the link between nonsymbolic and symbolic timing abilities in our dataset.

While we believe our temporal discrimination task tapped into children's nonsymbolic timing abilities, it is also possible that children may have relied upon some other quantitative dimension, such as the cumulative intensity of the stimuli, in order to make their responses. Although this remains a possibility, we do not believe this to be the case. While we predicted a relation between nonsymbolic and symbolic timing, there would be little reason to expect a relation between nonsymbolic cumulative intensity and symbolic timing. Thus, the strong correlation found between nonsymbolic and symbolic timing in the present study suggests that the temporal discrimination task is indeed tapping into children's timing abilities. Moreover, performance on our temporal discrimination task mimicked that of other studies investigating nonsymbolic timing using empty intervals, in which the duration to be timed is denoted by two brief tones at the beginning and end of the interval, in which cumulative intensity would not be an available cue (Grondin, 1993).

To fully investigate this possibility, future work should ask children to discriminate between two empty intervals.

Although we found a significant relation between nonsymbolic and symbolic timing abilities, our data cannot speak to whether this relation is causal or not, and if so, in which direction. One possibility is that learning symbolic temporal units of measurement may give children a more accurate understanding of the timing process. In this instance, after learning temporal symbols, children may be better able to encode durations and/or have heightened attention towards stimulus durations, leading to a more advanced ability to track time. Hints of a causal relation of this nature in the domain of number is supported by the work of Shusterman et al. (2016), in which it was found that young children learning to count demonstrate dramatic increases in nonsymbolic numerical acuity as soon as they master the basic counting principles. Similarly, Piazza et al. (2015) found adults in the Munduruku tribes of Brazil with zero years of formal schooling to have significantly less precise numerical acuity than individuals with one or more years of schooling, suggesting that formal education shapes nonsymbolic numerical abilities. However, whether the acquisition of timing symbols similarly leads to improvements in nonsymbolic timing abilities is an open question.

If it is the case that nonsymbolic timing is shaped by the acquisition of temporal symbols, what is the source of the improvement in timing? The acquisition of symbols may make temporal durations more salient and relevant to children, leading to heightened attention and temporal acuity. Alternatively, learning a system of measurement may provide children with strategies for tracking time, such as counting seconds. Although anecdotally, none of the children in our study overtly counted and the durations included in our temporal discrimination task were too short to make counting a less effective strategy, this possibility should be explored in future research.

Alternatively, the link between symbolic and nonsymbolic timing abilities may hint at the possibility that children with greater nonsymbolic temporal precision may have a better foundation upon which to learn temporal units of measurement. That is, it may be easier for children to learn temporal units of measurement when they already have a precise ability to track time and discriminate between durations. Again, evidence in the domain of number has suggested that nonsymbolic numerical abilities may form the foundation for the acquisition of numerical symbols. For example, nonsymbolic numerical discrimination abilities in infancy predict formal math achievement years later in preschool (Starr et al., 2013). Similarly, there is some evidence to suggest that training of nonsymbolic numerical abilities may even lead to improved math achievement in childhood (Hyde, Khanum, & Spelke, 2014; Park & Brannon, 2016; Wang, Odic, Halberda, & Feigenson, 2016). Future experimental work should investigate the nature of the relation between symbolic and nonsymbolic timing abilities to determine whether nonsymbolic timing abilities are predictive of symbolic acquisition and/or whether the acquisition of temporal symbols leads to subsequent improvements in nonsymbolic timing abilities.

If it is the case that a causal relation exists between symbolic and nonsymbolic timing, it would be important to explore why this may be and if other experiences may similarly shape timing abilities. For example, we know that nonsymbolic timing abilities become more precise over the course of development, even long before children acquire temporal symbols (Brannon et al., 2007; Feigenson, 2007; vanMarle & Wynn, 2006). Are there particular types of experiences that may promote temporal acuity more than others?

If so, is it possible to promote nonsymbolic timing in the service of acquiring more precise symbolic understandings? The answers to these questions are important for shedding light on timing processes while also having educational implications.

In conclusion, our data are the first to indicate that children's understandings of temporal units of measurement are correlated with individual differences in temporal acuity. Future work will be important for better characterizing this relation. Is this relation casual? If so, in which direction? Does the acquisition of symbolic timing sharpen nonsymbolic temporal acuity or does improving nonsymbolic temporal acuity promote the acquisition of temporal symbols? Or might a third variable contribute to this relation? Together, our study joins others in the field of numerical cognition suggesting that a child's ability to perceive quantity nonsymbolically is intricately linked to their formal understanding of symbols in that same quantitative domain.

Appendix. Temporal Units of Measurement Questionnaire

1. How many minutes are in an hour?
2. How many seconds are in a minute?
3. How many hours are in a day?
4. A. What is longer a minute or a second?

   B. What is longer a minute or an hour?

   C. What is longer a second or an hour?
5. A. What is longer a month or a day?

   B. What is longer a day or a week?

   C. What is longer a week or a year?

   D. What is longer a year or a month?
6. How many minutes does it take you to brush your teeth?
7. How many minutes does it take you to eat dinner?
8. How many minutes have you been talking to me?
9. Kristen jumped for a week, John jumped for a day. Who jumped longer?
10. Jennifer jumped for 5 seconds, Ben jumped for 2 minutes. Who jumped longer?
11. What does this clock say?
12. What time does this clock say?

13. Kit left home this morning at 9:05. Jane left home 20 minutes after Kit left. What time did Jane leave at?

CHAPTER 3: LEARNING ABOUT TIME DOES NOT REFINE CHILDREN'S TEMPORAL ACUITY

Abstract

Over development, children acquire symbols to represent abstract concepts such as time and number. Despite the importance of quantity symbols, it is unknown how acquiring these symbols impacts one's ability to perceive quantities (i.e., nonsymbolic representations). While it has been proposed that learning symbols shapes nonsymbolic quantitative abilities, this hypothesis has been understudied. Moreover, the research in support of this hypothesis has been solely correlational in nature, and thus, experimental manipulations are critical for determining whether this relation is causal. In the present study, kindergarteners and first graders (*N* = 133; who have yet to learn about temporal symbols in school) completed a temporal estimation task during which they were trained on (1) temporal symbols and effective timing strategies ("2 seconds" and counting on the beat), (2) training on temporal symbols only ("2 seconds"), or (3) a control training. Children's nonsymbolic and symbolic timing abilities were also assessed before and after training. Results revealed a strong correlation between children's nonsymbolic and symbolic timing abilities at pre-test (even when controlling for age), indicating this relation exists prior to formal classroom instruction on temporal symbols. Despite this, we found no support for the refinement hypothesis, as learning temporal strategies and/or symbols did not impact children's nonsymbolic timing abilities. Implications and future directions are discussed.

"How many blocks are you building with?" "How many minutes are in an hour?" Early education places a great emphasis on teaching young children units of measurement, or symbols (e.g., number words for number; minutes, seconds, hours…etc., for time), to represent quantities. During preschool and kindergarten, children's primary focus is learning numerical symbols (i.e., counting), and a shift to learning temporal symbols (i.e., minutes, seconds, hours…etc.) begins in early elementary school (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). Despite the ubiquity of quantity symbols, it is relatively unknown how knowledge of these symbols impacts one's ability to track these quantities (i.e., nonsymbolic representations). Does acquiring symbolic knowledge subsequently shape nonsymbolic abilities? The idea that formal symbolic knowledge may refine basic quantitative abilities – termed "the refinement hypothesis" (e.g., Mussolin, Nys, Content, & Leybaert, 2014; Suárez-Pellicioni & Booth, 2018) - has received some support from the numerical cognition literature, such that children's early math abilities and/or symbolic knowledge has been shown to predict their later nonsymbolic numerical abilities (Lyons, Bugden, Zheng, De Jesus, & Ansari, 2018; Matejko & Ansari, 2014; Mussolin et al., 2014; Shusterman et al., 2016). Notably, however, these investigations have been solely correlational in nature. In order to understand whether this relation is truly causal, experimental manipulations must be employed. In the present study, we conduct one of the first experimental investigations of the refinement hypothesis by testing whether learning temporal units of measurement – specifically acquiring knowledge about seconds – impacts children's ability to perceive time.

Evidence from number

A plethora of work has shown a relation between children's nonsymbolic numerical acuity and formal math abilities (i.e., symbolic abilities). This relation holds across all ages tested and across several math tasks (Bonny & Lourenco, 2013; Chen & Li, 2014; Libertus, Feigenson, & Halberda, 2011; Mazzocco, Feigenson, & Halberda, 2011; Odic, Lisboa, Eisinger, Olivera, Maiche, & Halberda, 2016; Schneider, Beeres, Coban, Merz, Schmidt, Stricker, & De Smedt, 2017). While researchers have promoted these findings as support for the idea that nonsymbolic numerical abilities form the foundation for the learning of formal symbolic number (e.g., Mazzocco et al., 2011; Starr, Libertus, & Brannon, 2013), it is also possible that the acquisition of formal symbols refines nonsymbolic numerical abilities. Evidence in favor of this hypothesis reveals 3-4 year olds' performance on a symbolic number battery (e.g., counting, cardinality, etc.) predicted nonsymbolic number performance seven months later, whereas nonsymbolic number performance did not predict later symbolic abilities (Mussolin, Nys, Content, & Leybaert, 2014; see also Lyons, Bugden, Zheng, De Jesus, & Ansari, 2018). In a related study, Matejko and Ansari (2014) found both symbolic and nonsymbolic performance at the beginning of the school year predicted later nonsymbolic performance, whereas earlier symbolic performance was the only significant predictor of symbolic performance. Relatedly, Shusterman et al. (2016) discovered that as soon as preschoolers acquired an understanding of cardinality (the last number counted represents the size of the set), they demonstrated enhanced nonsymbolic numerical discrimination. Neuroimaging data also support the refinement hypothesis: math performance predicted bilateral IPS activity during a numerical discrimination task two years later, but not in the other direction (Suárez-Pellicioni & Booth, 2018).

Although these cross-lagged, correlational studies provide support for the refinement hypothesis by testing children's symbolic and nonsymbolic abilities at two or more time points, only a single study has experimentally manipulated symbolic knowledge in order to test possible changes in nonsymbolic abilities. In this work, adults were explicitly trained on an arithmetic task (outside of the traditional classroom experience) and their numerical acuity was tested before and after training (Lindskog, Winman, & Poom, 2016). While the arithmetic training did not impact subsequent numerical acuity, the study was conducted with adults, who already have a substantial understanding of symbolic number. Given children's limited symbolic knowledge, and the malleability of children's nonsymbolic abilities (see Wang, Odic, Halberda, & Feigenson, 2016), determining whether explicit instruction of symbols would enhance children's nonsymbolic acuity may be a stronger test of the refinement hypothesis.

What about time?

While the field of numerical cognition has been inundated with investigations of this relation, little work has investigated similar relations in other quantitative domains, such as time. Mirroring the numerical literature, recent work found a correlation between 6-7 year olds' ability to discriminate durations (i.e., their nonsymbolic representations of time) and their understanding of temporal units of measurement (i.e., symbolic representations of time; Hamamouche & Cordes, accepted). This study tested 6-7 year olds because that is the age at which children typically begin receiving formal instruction on temporal units of measurement in the classroom. However, it is an open question whether this relation holds in children who have yet to receive formal instruction on temporal units of measurement. Moreover, while a correlation has been established, it is unclear whether this relation is causal, and specifically, whether the refinement hypothesis holds in the domain of time. That is, it is unknown whether acquiring temporal units of measurement subsequently shapes children's nonsymbolic timing abilities. The relation in the domain of number may be facilitated as both numerical symbols and nonsymbolic abilities are discrete. It may be more difficult, however, for discrete symbols to shape continuous, nonsymbolic quantities, as is the case for time. Thus, in the current work, we test 1) whether a correlation exists between children's symbolic and nonsymbolic abilities prior to formal instruction on time, and 2) if learning temporal units of measurement results in enhanced nonsymbolic timing abilities.

The Current Study

In the current study, we investigate the role that learning temporal units of measurement plays in children's perception of time. We specifically chose to study kindergarten and first semester first graders who have yet to receive extensive instruction on temporal units of measurement in the classroom (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010) and whose nonsymbolic temporal acuity is still improving (Odic, 2018). In order to determine whether acquiring symbolic knowledge impacted children's nonsymbolic representations of time, children participated in one of three symbolic training conditions: (1) in the Strategies condition, children were taught to keep track of seconds by counting on a beat; (2) in the Symbols condition, children only received instruction on temporal symbols without information about how to track the time; or (3) in a Control condition, in which children received no instruction on temporal symbols. Symbolic and nonsymbolic timing abilities were measured before and after training. Children also completed a nonsymbolic numerical task, in order to test whether acquiring temporal symbols uniquely impacted children's temporal acuity, or impacts nonsymbolic quantitative abilities more generally. We predicted that children in both the Strategies and Symbols condition should have enhanced temporal discrimination task after training. However, those in the Strategies condition may show even greater improvements in nonsymbolic abilities.

Methods

Participants

Kindergarten and first grade children (*N* =133; M_{age} = 76.37 months, 69 females) were recruited from the greater Boston (*n* = 44) and greater Indianapolis area (*n* = 89). All children were tested in their schools during the school day. One additional participant completed the study in the lab. Required sample size was determined on a priori calculations in G*Power (Faul et al., 2009) requiring 0.80 power for detecting a medium effect size on central analyses.

An additional 33 participants were excluded for the following reasons: 18 participants did not have a complete data set (completing the wrong estimation task: *n* = 1, opting out of the study: *n* = 6, missing data due to computer error: *n* = 5, or only completing the first visit due to scheduling restraints: *n* = 6); 7 participants were highly distracted during the training task; 3 participants did not understand the training task; and 5 performed below chance at both pre- and post-test.

Procedure

Children completed two test sessions which occurred within seven days. Children completed tasks in a fixed order. At pre-test, children completed 1) temporal discrimination (an assessment of nonsymbolic timing abilities), 2) two temporal estimation tasks (an assessment symbolic mapping; order counterbalanced), 3) numerical discrimination (assessing nonsymbolic numerical abilities), 4) the temporal units of measurement questionnaire (assessing symbolic abilities), and 5) training (dependent upon condition). At post-test, children began with the training task, and then completed the pre-test measures in the same order. Children did not complete the temporal units of measurement questionnaire at post-test.

(1) Temporal discrimination (Hamamouche & Cordes, 2019)

The temporal discrimination task was a measure of children's nonsymbolic timing abilities. Children saw two animals (a chicken and a cow) on the computer screen and each animal took turns playing a musical instrument for a certain amount of time. The chicken always played first followed by the cow. While the animal played, children heard a sound (horn for chicken, train for cow) and saw a box appear around the animal (i.e., both visual and auditory stimuli). Children were asked to indicate which animal played their instrument longer on every trial. Children completed two practice trials (involving a 3-fold difference in the two durations), during which they could ask the experimenter questions about the task. After the practice trials, children put on headphones so they could clearly hear the stimuli. Because temporal strategies are only useful during supra-second timing (Grondin, Meilleur-Wells, & Lachance, 1999), we predicted that timing supra-second durations (durations > 1000ms) may be more susceptible to modification following the learning temporal symbols and strategies compared to sub-second duration timing (durations < 1000ms). Thus, half of the trials involved durations in the sub-second range (< 1000ms) and half involved durations in the supra-second range (> 1000ms). Moreover, to vary the difficulty of the task, pairs of durations presented on every trial differed by one of three ratios: 1.25, 1.5, and 2. Children completed 18 test trials (2 duration ranges x 3 ratios x 3 times each).

(2) Temporal estimation

Children completed two distinct temporal estimation tasks (order counterbalanced) to test the mapping between nonsymbolic and symbolic representations of time. We included two different tasks as they tapped into different directions of mapping. By doing so, we hoped to explore bidirectional mappings between nonsymbolic and symbolic representations of time and determine whether either direction of mapping was more accessible for young children.

In the symbolic to nonsymbolic temporal estimation task, children were asked to hold a specified key for “three seconds”, “seven seconds”, and “nine seconds” (order counterbalanced). A blue oval appeared on the screen while children pressed the key. In the nonsymbolic to symbolic temporal estimation task, children saw a blue oval displayed for a certain duration (6 seconds at pre-test and 4 seconds at post-test) and were asked to verbally estimate how many seconds the oval was on the screen.

(3) Numerical discrimination Task (Panamath; Halberda et al., 2008)

On every trial of the numerical discrimination task, children saw two arrays of dots (one blue and one yellow) and were asked to identify which array contained more dots. The arrays were displayed briefly (1951 ms), so that children were unable to count the number of dots in each array. Because the Panamath program (Halberda et al., 2008) adjusts the difficulty and number of trials based on the age of the participants, we set the age parameter to 5 and the test duration to 1 minute for all participants.

(4) Temporal symbols questionnaire (modeled after Hamamouche & Cordes, accepted)

The temporal symbols questionnaire consisted of nine questions regarding children's symbolic understanding of time. The experimenter read the questions aloud to the child and recorded their responses. Questions included asking factual information about temporal units (e.g., How many minutes are in an hour?), deciding which of two temporal units of measurement was longer (e.g., What is longer a minute or a second?), and estimating how long it took to do everyday activities (e.g., eat dinner, brush teeth).

(5) Training

Children were randomly assigned to one of three training conditions: Strategies (n = 45, M_{age} = 76.22 months), Symbols (n = 44, M_{age} = 76.32 months), and Control (n = 44, M_{age} = 76.57 months). The training task was identical on both testing days.

Strategies Condition. To introduce children to the task, the experimenter told children they were going to learn about time, and specifically see what 10 seconds felt like. To do so, the experimenter held a timer, and counted by saying One Mississippi, Two Mississippi,…etc. as each second passed to introduce a counting strategy to the children. After one example, the experimenter explained that it was easier to keep track of the seconds by counting them on a specific beat. Then, the experimenter repeated this process and encouraged children to count the seconds with them using the new Mississippi strategy. After this introduction to timing, children began the training task in which they were told that they would be playing hide and seek with a cat on a computer screen. In order to play the game, children would watch the cat on the screen, and when it disappeared, children would press the space bar for the same amount of time the cat had been on the screen. At the beginning of each trial, the experimenter stated how long the cat would appear on the screen (e.g., "The cat is going to be there for 2 seconds."). After the cat disappeared, the experimenter prompted children by saying e.g., “That was 2 seconds. Can you press the space bar for 2 seconds?” Trial-by-trial feedback was provided. When a response was within 20% of the correct response (e.g., responses between 4-6 seconds were considered correct when asked to produce 5 seconds), a smiley face appeared on the screen for 5000ms. If the response was outside of the 20% range, a frowning face appeared on the screen for 5000ms, with feedback “That was too long [short]. Try to hold the space bar for shorter [longer] next time.” At both pre- and post-test, children completed 10 practice trials (5 trials involving 2 seconds followed by 5 trials involving 3 seconds) followed by 15 test trials (3 trials using 2 seconds, then 3 trials using 3 seconds,…continuing through 6 seconds). To increase the amount of exposure to the counting strategy, the experimenter counted the seconds as the cat was on the screen during the practice trials. During the test trials, if children did not explicitly count while pressing the space bar, the experimenter reminded children to count the seconds to help them. Moreover, when children did not count on two trials in a row, the experimenter counted the seconds that the cat appeared on the screen for the next trial.

Symbols Condition. The symbols condition was identical to the temporal strategies condition with the following exceptions. The experimenter told the child they would learn about 10 seconds and showed children a timer and allowed 10 seconds to pass; however, at no point during the training did the experimenter provide children with any strategies to help them, nor did they encourage children to use counting strategies to complete the training task. The prompts before and after each training trial were identical to the temporal strategies condition. Again, trial-by-trial feedback was provided.

Control Condition. Children assigned to the control condition did not receive the introduction to the task using the timer. Moreover, before each trial, the experimenter alerted children to the start of the trial by saying, “The cat is going to be there.” After each trial, the experimenter prompted the child by saying, “Can you press the space bar for the same amount of time?” Children continued to receive trial-by-trial feedback.

On the second day of testing, children first completed the training (which was identical to pre-test training) and then subsequently completed the temporal discrimination task, temporal estimation task (order counterbalanced, in the same order as pre-test), and the numerical discrimination task. At the end of both sessions, children received two stickers for participating.

Data Coding

Temporal discrimination

Proportion correct served as the dependent variable for the temporal discrimination task. We also calculated the proportion correct for performance in the sub- and supra-second ranges, as well as, for each individual ratio. Only one participant had an overall temporal discrimination performance that was an outlier (Pre-Test, performance three standard deviations below the mean) and thus was excluded from analyses involving the temporal discrimination task.

Temporal estimation: symbolic to nonsymbolic

Average estimate and slope were the dependent variables for the symbolic to nonsymbolic temporal estimation task. The measure of average estimate allows us to determine whether participants are over- or underestimating, whereas the slope analysis indicates whether participants are adjusting their estimates based on the duration of the stimulus. Data points three standard deviations above/below the mean were excluded from analyses (Average Estimate: Pre-Test: n_{Above} = 3; Post-test: n_{Above} = 3; Slope: Pre-Test: n_{Above} = 2, n_{Below} = 1; Post-Test: n_{Above} = 1, n_{Bleow} = 2).

Temporal estimation: nonsymbolic to symbolic

Because the duration presented differed from pre-test to post-test, we calculated absolute average error as the dependent variable for the nonsymbolic to symbolic temporal estimation task. We first removed any responses that were non-numerical (i.e., "a little", "circle"; Pre-test: n = 3, Post-test: n = 1). Then, we excluded any values that were three standard deviations above/below the mean. Two participants at pre-test and another at post-test had responses three standard deviations above the mean, and were thus excluded from analyses involving this estimation task.

Numerical Discrimination

Percent correct was the dependent variable for the numerical discrimination task, as this has been shown to be the most reflective of performance (Inglis & Gilmore, 2014). Data points three standard deviations above or below the mean (Pre-test: n_{Below} = 3, Post-test: n_{Below} = 3) were excluded from analyses involving this task.

Temporal symbols questionnaire

Performance on the temporal symbols questionnaire was scored for accuracy. Instances in which the child refused to provide a response, or indicated that they did not know the answer, were scored as incorrect. There were no outliers.

Training

Accuracy was recorded during the test trials of the training task. Given predicted differences in accuracy across the conditions (performance should be most accurate in the strategies condition, followed by the symbols, and lastly the control condition), we did not exclude participants based on their performance on the training task.

Results

Average performance on each task at pre- and post-test can be found in Table 3.1, and correlations between pre- and post-test performance can be found in Table 3.2. There were no differences in performance at pre-test across our two research sites (p's >.15) or conditions (p's > .3). Pre-test performance on all tasks (p's < .01) was found to be correlated with age; except the nonsymbolic to symbolic estimation and symbolic to nonsymbolic (slope) p's > .3; thus, analyses included age as a covariate to control for potential age differences.

	Temporal Discrimination	Temporal Estimation (Nonsymbolic – Symbolic) Average Error	Temporal Estimation (Symbolic – Nonsymbolic) Average Estimate	Temporal Estimation (Symbolic – Nonsymbolic) Slope	Temporal Symbols Questionnaire	Numerical Discrimination	Training
Strategies							
Pre-Test	.75 (.01)	.77 (.07)	11.73 (2.21)	1.30 (.40)	.56 (.04)	92.05 (1.29)	.61 (.04)
Post-Test	.72 (.02)	.35 (.07)	6.02 (.41)	.92 (.09)	--	91.90 (1.31)	.68 (.04)
Symbols							
Pre-Test	.74 (.02)	.80 (.08)	8.77 (1.70)	1.05 (.26)	.52 (.04)	93.46 (1.26)	.44 (.04)
Post-Test	.69 (.02)	.84 (.09)	4.30 (.41)	.62 (.07)	--	91.28 (1.14)	.45 (.04)
Control							
Pre-Test	.73 (.02)	.67 (.08)	12.46 (1.93)	1.38 (.45)	.52 (.04)	91.68 (1.35)	.42 (.04)
Post-Test	.70 (.02)	.88 (.10)	3.47 (.26)	.44 (.05)	--	92.73 (1.31	.44 (.03)

Table 3.1. Average performance (SE) on each task at pre- and post-test.

	Temporal Discrimination	Temporal Estimation (Nonsymbolic – Symbolic) Average Error	Temporal Estimation (Symbolic – Nonsymbolic) Average Estimate	Temporal Estimation (Symbolic – Nonsymbolic) Slope	Numerical Discrimination
Strategies	.357*	-.023	.489**	-.239	.522**
Symbols	.370*	.573**	.186	.442**	.423*
Control	.276^{+}	.551**	.383*	-.041	.510*

Table 3.2 Correlations between pre- and post-test performance.
***$p < .001$, *$p < .05$, $^{+}p < .1$

Correlation between Nonsymbolic and Symbolic Timing. We first tested for a correlation between nonsymbolic and symbolic timing abilities (as measured by the temporal units of measurement questionnaire) at pre-test. This correlation held when controlling for age, $r = .212$, $p = .015$, See Figure 3.1.

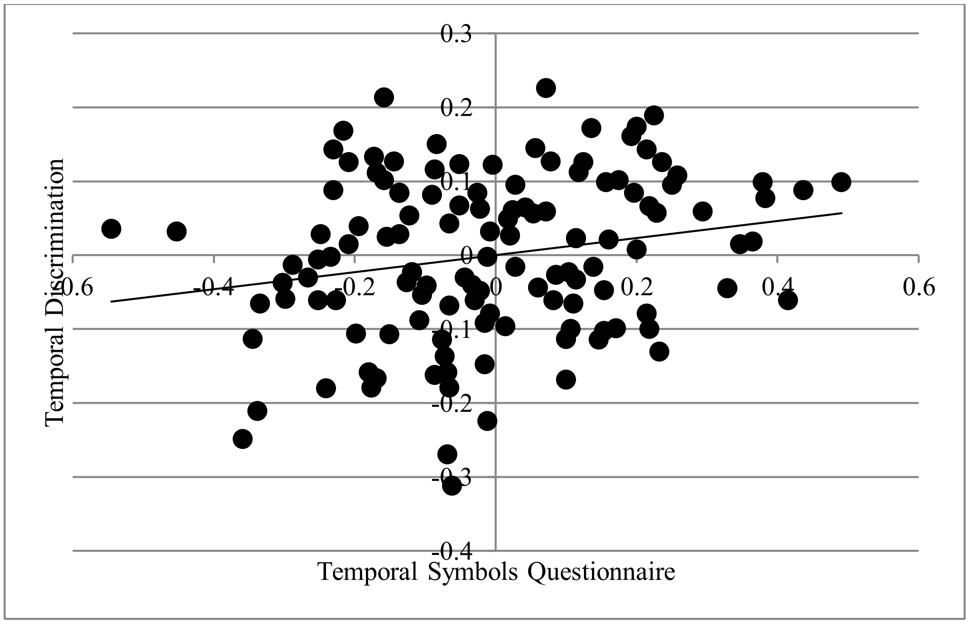

Figure 3.1. There was a strong correlation between temporal discrimination performance and an understanding of temporal units of measurement when controlling for age.

Effect of Training

We next investigated whether accuracy on the training task differed amongst the conditions. To do so, we conducted a univariate ANOVA exploring the effect of Condition (as a fixed factor) and Age (as a covariate) on accuracy on the training task. Both Condition, $F(2, 129) = 16.620$, $p < .001$, and Age ($F(1, 129) = 21.590$, $p < .001$) predicted training performance. Parameter estimates revealed that with each additional month of age, children responded correctly on an additional half item (B = .011). Moreover, children in the Strategies ($M = .646$, $SE = .030$) condition outperformed those in both the Symbols ($M =$

.447, *SE* = .030, $p < .001$) and the Control ($M = .428$, $SE = .030$, $p < .001$) conditions. There was no significant difference between those in the Symbols or Control conditions ($p = .659$), See Figure 3.2.

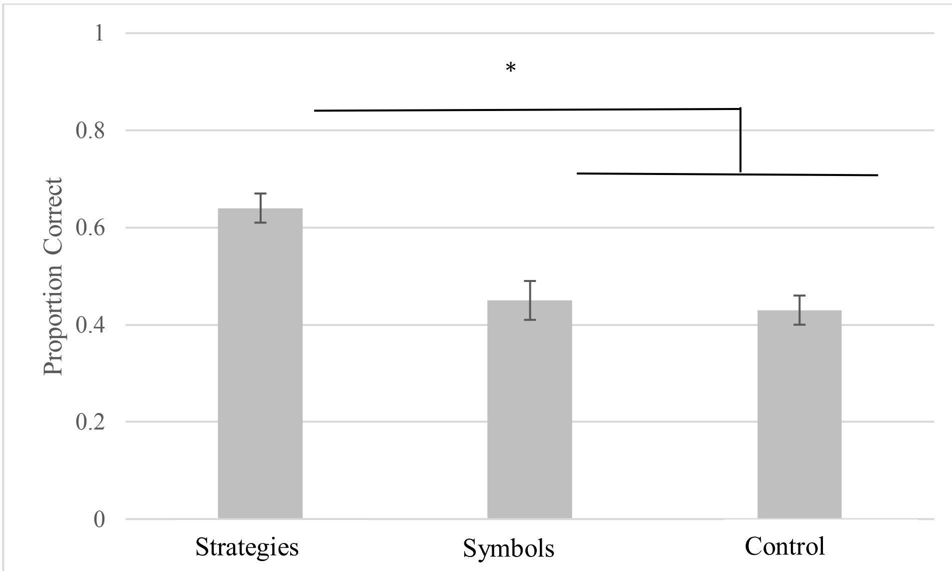

Figure 3.2. Children who learned strategies outperformed their peers on the training task.

Temporal Discrimination

Performance was above chance at both pre-test and post-test on the temporal discrimination (Pre-Test: $M = .74$, $SE = .01$; $t(131) = 24.942$, $p < .001$; Post-Test: $M = .71$, $SE = .01$; $t(132) = 18.403$, $p < .001$). Moreover, a 2 (Time: Pre- versus Post-test) x 3 (Ratio: 1.25, 1.5, 2) repeated measures ANOVA on temporal discrimination performance revealed a main effect of Time, $F(1, 131) = 7.564$, $p = .007$, $\eta_p^2 = .055$, such that performance was worse at post-test ($M = .707$, $SE = .011$) compared to pre-test ($M = .741$, $SE = .010$). There was also a main effect of Ratio, $F(2, 262) = 88.270$, $p < .001$, $\eta_p^2 = .403$, such that performance increased as the ratio between the two durations increased. The Time x Ratio interaction did not reach significance, $p > .5$.

In order to determine whether the training directly impacted children's temporal discrimination performance, we first conducted a 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA with age as a covariate on children's performance on the nonsymbolic temporal discrimination task. Age was related to discrimination performance, $F(1, 128) = 14.170$, $p < .001$, $\eta_p^2 = .100$. Parameter estimates revealed that with each additional month, children responded correctly on approximately one additional trial (Pre-Test: B = .004; Post-Test: B = .005). No other main effects or interactions emerged, *p's* > .2, suggesting that our training conditions did not have a differential impact on temporal discrimination performance, See Figure 3.3.

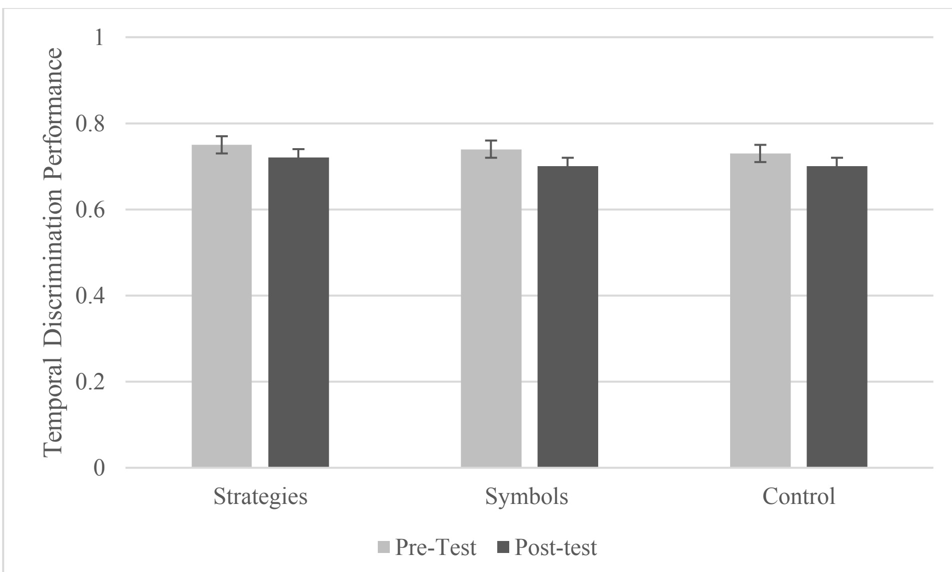

Figure 3.3. Regardless of Condition, temporal discrimination performance was comparable at pre- and post-test.

Sub- and Supra-Second Temporal Discrimination. A secondary question concerned whether temporal precision may be differentially impacted depending on the range of durations tested. That is, we reasoned that our training may specifically impact durations in the supra-second range (compared to the sub-second range) as supra-second duration timing may be more effectively measured via counting strategies (it is hard to count more than one beat for durations less than a second) and because our symbols specifically trained children on supra-second durations (i.e., children were taught about seconds, not about milliseconds). To investigate this possibility, we performed identical repeated measures ANCOVAs using sub-second and supra-second performance (separately, again controlling for age). Both ANCOVAs revealed a main effect of Age, (sub-second: $F(1, 128) = 8.058$, $p = .005$, $\eta_p^2 = .059$; supra-second: $F(1, 128) = 11.083$, $p = .001$, $\eta_p^2 = .080$) and no other main effects or interactions reached significance, *p's* > .2.

Temporal Estimation (Symbolic to Nonsymbolic)

Average Estimate. Children's average estimates was greater than the true mean ($M = 6.34$) at pre-test, but smaller than the true mean at post-test (Pre-Test: $M = 10.97$, $SE = 1.13$; $t(129) = 4.10$, $p < .001$; Post-Test: $M = 4.58$, $SE = .23$, $t(129) = -7.657$, $p < .001$).

We ran a comparable 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA on children's performance on the symbolic to nonsymbolic temporal estimation task with age as a covariate. There was a main effect of Time, $F(1, 124) = 11.030$, $p = .001$, $\eta_p^2 = .082$, such that the average estimate was smaller at post-test ($M = 4.62$, $SE = .215$) compared to pre-test ($M = 10.38$, $SE = 1.037$). There was also a main effect of Age, $F(1, 124) = 6.134$, $p = .015$, $\eta_p^2 = .047$, that was qualified by a Time x Age interaction, $F(1, 124) = 7.748$, $p = .006$, $\eta_p^2 = .059$. To follow-up on this interaction, we first looked for a correlation between age and change in estimates. A significant correlation emerged, $r = .236$, $p = .007$. Plotting this correlation indicated that younger children substantially lowered their estimates from pre-test to post-test, whereas older children did not change their estimates as much, See Figure 3.4. Finally, there was a marginal Time x Condition interaction, $F(2, 124) = 2.817$, $p = .064$, $\eta_p^2 = .043$. Children in all conditions produced significantly shorter estimates at post-test compared to pre-test (Strategies: Pre-Test: M = 9.91, SE = 1.88; Post-Test: M = 6.02, SE = .42; p = .029; Symbols: Pre-Test: M = 8.77, SE = 1.7; Post-Test: M = 4.30, SE = .41; p = .010; Control: Pre-Test: M = 12.46, SE = 1.93; Post-Test: M = 3.45, SE = .27; $p < .001$). To further explore this interaction, we conducted a univariate ANCOVA with Age as the covariate and Condition as the fixed factor predicting the difference in average estimate. Results revealed that the change in estimates in the Control condition was significantly greater than the change in the Strategies condition (p = .034) and marginally greater from those in the Symbols condition (p = .053). The Strategies and Symbols condition showed a similar change in estimates over time (p = .820). See Figure 3.5. The main effect of Condition did not reach significance, p = .454.

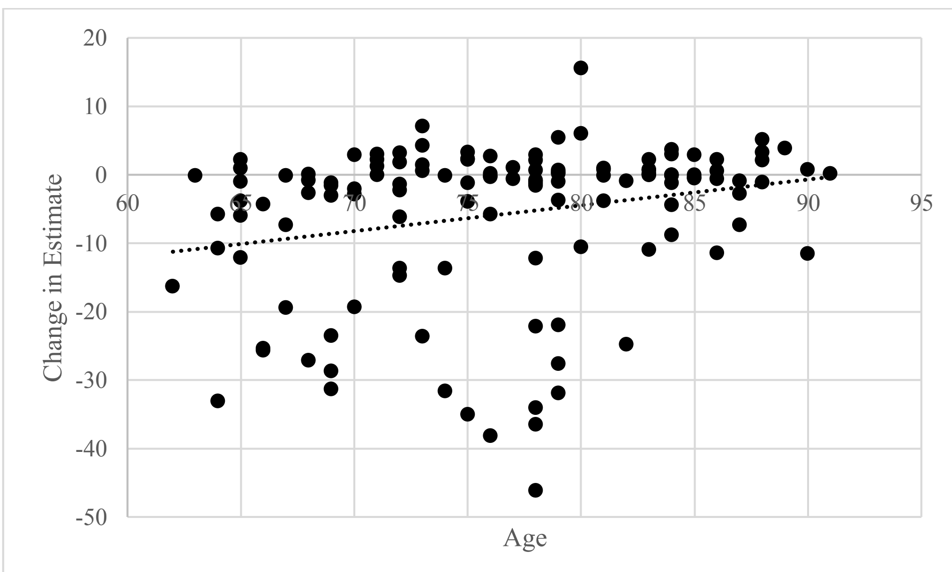

Figure 3.4. Older children maintained comparable estimates over time, whereas younger children provided shorter estimates.

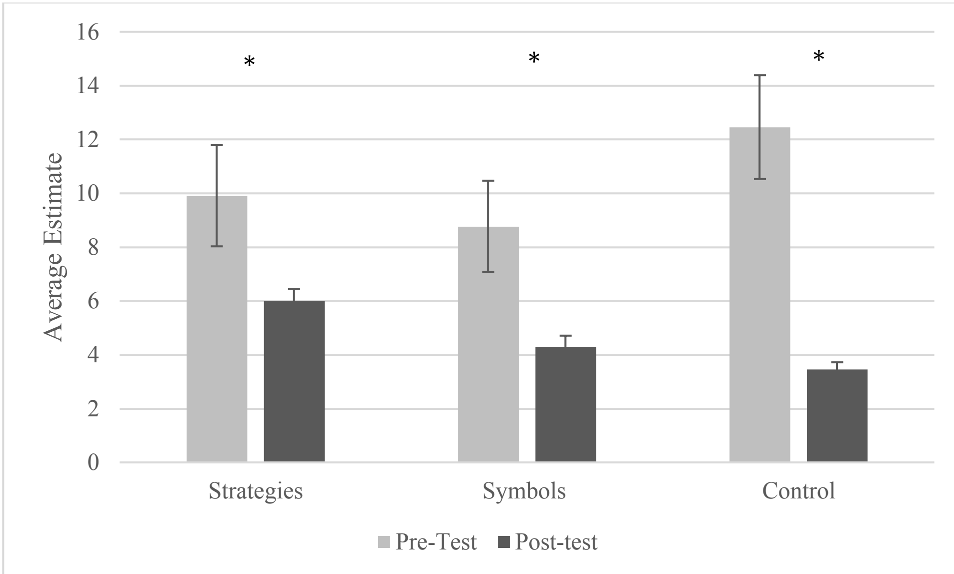

Figure 3.5. Children's estimates were significantly shorter at post-test for children in all conditions.

Next, we conducted one sample t-test to determine whether the average estimate at post-test was different from the true average in each condition separately. Results revealed that those in the Strategies condition had accurate estimates at post-test ($M = 6.02$, $SE = .41$, $t(41) = -.778$, $p = .441$), whereas those in the Symbols and Control condition significantly underestimated at post-test (Symbols: $M = 4.30$, $SE = .21$, $t(43) = -4.937$, $p < .001$; Control: $M = 3.47$, $SE = .26$, $t(43) = -10.994$, $p < .001$). Thus, children who acquired timing strategies were more accurate after training.

Slope. A one sample t-test revealed that children's slopes were significantly greater than 0, indicating that estimates increased as the duration to be estimated increased at pre- and post-test (Pre-Test: $M = .66$, $SE = .06$; $t(89) = 10.225$, $p < .001$; Post-Test: $M = .66$, $SE = .04$; $t(129) = 14.775$, $p < .001$).

An identical 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA on response slopes relating symbolic to nonsymbolic temporal estimation was conducted, with age as a covariate. No main effects or interactions reached significance, *p's* > .1.

To determine whether performance became more precise at post-test (as indicative of a slope closer to 1), we performed one-sample t-tests on post-test slope for each condition separately. While performance in the Strategies condition was very precise and did not differ from 1 (M = .92, SE = .09, $t(42)$ = -.935, p = .355), those in the Symbols and Control conditions had slopes significantly less than 1 (Symbols: M = .62, SE = .07; $t(42)$ = -5.331, p < .001; Control: M = .44, SE = .05; $t(43)$ = -10.231, p < .001). See Figure 3.6.

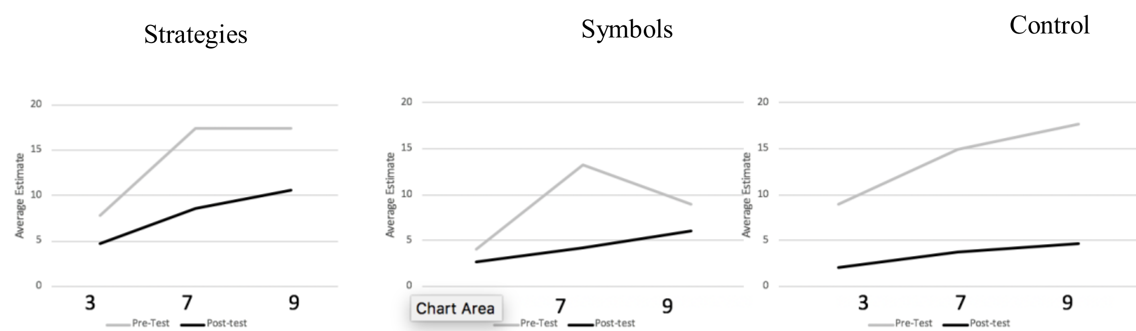

Figure 3.6. The slope indicated more precise performance at post-test for those in the Strategies condition.

Temporal Estimation (Nonsymbolic to Symbolic)

<u>Absolute Error</u>. A 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA on children's performance on the symbolic to nonsymbolic temporal estimation task (with Age as a covariate) revealed a main effect of Condition, $F(2, 122) = 1.673$, $p = .024$, $\eta_p^2 = .059$, that was qualified by a Time x Condition interaction, $F(2, 122) = 14.318$, $p < .001$, $\eta_p^2 = .190$. Children assigned to the Strategies (Pre-Test: M = .77, SE = .07; Post-Test: M = .34, SE = .07; $t(41) = 4.123$, p <

.001) produced responses with significantly less error at post-test compared to pre-test, whereas those in the Control condition produced more error at post-test (Pre-Test: $M = .63$, $SE = .08$; Post-Test: $M = .90$, $SE = .11$; $t(41) = -2.93$, $p = .006$). Those in the Symbols condition showed no improvement after training (Pre-Test: $M = .80$, $SE = .09$; Post-Test: $M = .84$, $SE = .09$; $t(41) = -.456$, $p = .651$). Moreover, while children in the Strategies and See Figure 3.7. No other main effects or interactions reached significance, *p's* > .3.

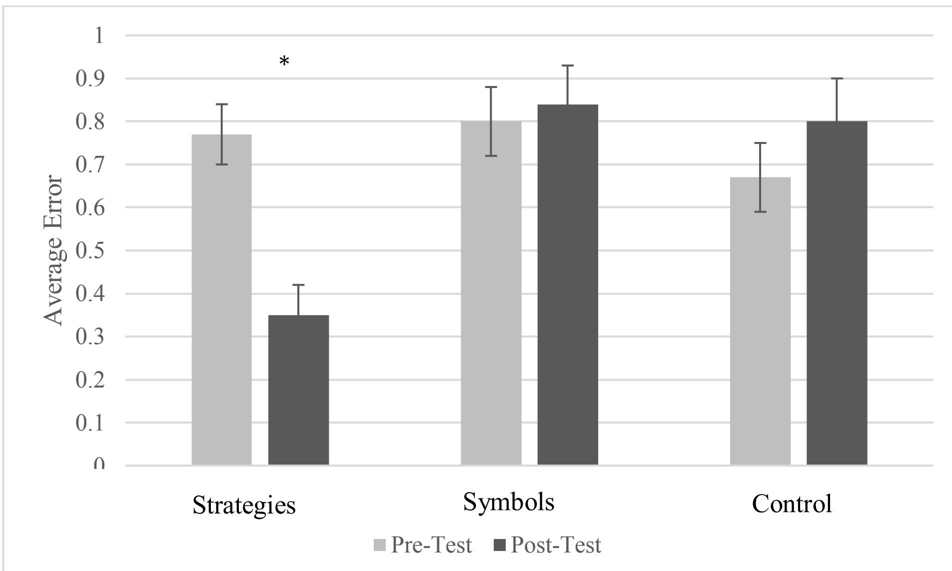

Figure 3.7. Error decreased significantly in the Strategies condition but not in the Symbols or Control conditions.

Numerical Discrimination.

Numerical discrimination performance was above chance at both time points (Pre-Test: $M = 92.45$, $SE = .75$; $t(129) = 56.61$, $p < .001$; Post-Test: $M = 91.97$, $SE = .72$; $t(129) = 58.044$, $p < .001$). A 2 (Time: Pre- versus Post-Test) x 4 (Ratio: 1.28, 1.47, 1.75, 2.75) repeated measures ANOVA revealed a main effect of Ratio, $F(3, 381) = 36.475$, $p < .001$, $\eta_p^2 = .223$, such that performance improved as the ratio between the two numbers increased. No other main effects or interactions reached significance, *p's* > .4.

Finally, to test whether learning about temporal units of measurement uniquely impacted nonsymbolic timing or nonsymbolic abilities more generally, we conducted a comparable 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA on children's performance on the numerical discrimination task, including Age as a covariate. There was a main effect of Age, $F(1, 124) = 12.664$, $p = .001$, $\eta_p^2 = .093$. With each increasing month of age, children answered ~5 more trials correct at pre-test (B = .310) and ~4 more trials correct at post-test (B = .268) better at post-test. No other main effects or interactions reached significance, *p's* > .2.

Discussion

We began this work with two questions: 1) Is there a correlation between children's understanding of temporal units of measurement and nonsymbolic timing prior to formal classroom instruction on temporal symbols? 2) Does learning temporal symbols enhance subsequent nonsymbolic timing? By training children on temporal units of measurement, we were able to perform the first experimental investigation of the refinement hypothesis in children in the domain of time.

Correlation between Nonsymbolic and Symbolic Timing

While a correlation has been found between 6-7 year olds' understanding of temporal units of measurement and nonsymbolic timing (Hamamouche & Cordes, accepted), it was previously unknown whether this correlation holds prior to formal instruction on temporal symbols. Our results replicate and extend these findings to show this correlation is present in children who have yet to receive formal instruction on timing. Importantly, this correlation was strong even when controlling for age, indicating that this relation is not due to age-related changes in temporal acuity, but instead reflects a unique relation between symbolic and nonsymbolic timing abilities. Joining previous work, this finding highlights the strength of the relation between symbolic and nonsymbolic timing both before and during the acquisition of temporal units of measurement in the classroom.

The Refinement Hypothesis

Our second aim was to test the refinement hypothesis by teaching children about temporal units of measurement and subsequently testing their nonsymbolic timing abilities. Importantly, the training task was identical across all conditions, with a few exceptions. Children in the Strategies condition learned a timing strategy (specifically counting on the beat). The experimenter also labeled the duration before and after each trial for children in the Strategies and Symbols conditions. Despite the similarity in training, results revealed that children who were introduced to the timing strategy outperformed their peers on the training task, indicating that our training was successful. Although we hypothesized that both the Strategies and Symbols conditions should help children acquire symbolic knowledge of timing, this was not the case – at least based on training performance only. As predicted, children in the Strategies condition outperformed their peers in both the Symbols and Control conditions. Unexpectedly though, children in the Symbols and Control conditions performed comparably on the training. Thus, counter to our predictions, being exposed to temporal symbols was no more beneficial than receiving no instruction on timing.

Results of the temporal estimation tasks provide further evidence that the training impacted children's timing abilities. Children in the Strategies condition produced less error at post-test compared to pre-test, whereas children in the Control condition produced more error at post-test, and those in the Symbols condition produced comparable error at both time points. These results suggest that children who were exposed to strategies did indeed learn about timing and were able to transfer this learning to a different temporal estimation task. Moreover, analyses of average estimate and slope reflected precise performance revealed that children in the strategies condition also showed greater precision in the symbolic to nonsymbolic temporal estimation task. Thus, as predicted, acquiring temporal strategies showed benefits above and beyond solely learning temporal symbols.

Although previous work in the domain of numerical cognition has provided some support for the refinement hypothesis, such that acquiring numerical symbols shapes numerical acuity (e.g., Lyons et al., 2018; Matejko & Ansari, 2014; Mussolin et al., 2014), only a single study has tested this experimentally. Moreover, mapping discrete temporal symbols on continuous nonsymbolic representations of time may be more difficult than mapping discrete numerical symbols onto discrete nonsymbolic numerical representations. As such, testing the effect of learning temporal symbols on temporal acuity may be particularly beneficial for determining the importance of structural isomorphisms within the refinement hypothesis. Despite the effectiveness of our training, we found no evidence of refinement. Regardless of training condition, children's nonsymbolic timing abilities remained consistent from pre- to post-test. Given the present findings, it is possible that structural similarities may be important for the refinement hypothesis. However, it is important to note that our results mirror Lindskog (2016) in which training adults' arithmetic abilities did not enhance numerical acuity. Thus, although correlational data hint at the possibility of the refinement hypothesis, experimental manipulations do not support this possibility. If there continues to be little support for the refinement hypothesis, the cause of improvements in nonsymbolic acuity remains unclear. While lack of support for the refinement hypothesis discounts the possibility that symbol acquisition enhances nonsymbolic representations, general maturation (e.g., Lipton & Spelke, 2003) and/or language acquisition (e.g., Posid & Cordes, 2015) may still play a role in improving nonsymbolic abilities. Future research will be critical for understanding the cause of improving nonsymbolic abilities with development.

Although we did not find evidence for the refinement hypothesis, it is important to note that our training did not reflect formal classroom instruction on temporal units of measurement. While children typically spend months learning about symbols in the classroom, our training took place over the course of two days and lasted approximately 20 minutes total. Additionally, our training focused on learning about seconds, whereas early instruction on temporal symbols is typically focused on understanding hours and minutes and telling time. A training that more closely reflects traditional classroom instruction on temporal units of measurement (i.e., over the course of several weeks) may serve as a stronger test of the refinement hypothesis. Although not involving an experimental manipulation, it may also be fruitful to test children's temporal discrimination abilities at the beginning and end of 1st grade to determine whether gains in nonsymbolic abilities correlates with formal instruction on temporal units of measurement.

Conclusions

In conclusion, the current study replicated the relation between symbolic and nonsymbolic timing in a younger group of children. Critically, these children had yet to learn formal temporal symbols in school, suggesting that this relation is not contingent upon formal symbol acquisition in school. Although this was the first experimental manipulation testing the refinement hypothesis in timing, we found learning temporal symbols did not impact subsequent nonsymbolic timing abilities. Given the limited amount of work that has been done thus far, exploring the refinement hypothesis remains a rich area of future investigation.

Appendix. Covariate Analyses

We also conducted all repeated measures ANCOVAs including pre-test performance as a covariate. By doing so, we are able to test how pre-test performance interacts with Time and Condition. However, including pre-test performance as a covariate in the model adjusts all pre-test performance to be identical, and thus makes it difficult to compare Time x Condition interactions as pre-test performance is adjusted and not reflective of the true means in each condition. In the following analyses, we highlight interactions between our covariates (age and pre-test performance) and Time and Condition separately.

Temporal Discrimination

The 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA (including age and pre-test temporal discrimination performance as covariates) on temporal discrimination performance revealed a Time x Age interaction, $F(1, 127) = 6.944$, $p = .009$, $\eta_p^2 = .052$. Although there was no significant correlation between age and change in discrimination performance ($r = .081$, $p = .354$), plotting the relation between the variables indicated that on average younger children tended to show worse performance at post-test, whereas older children showed no change, See Figure A3.1. There was also a Time x Pre-Test performance interaction, $F(1, 127) = 47.881$, $p < .001$, $\eta_p^2 = .274$. Plotting this relation indicated that children who performed poorly at pre-test tended to have improved performance at post-test, whereas children who performed better at pre-test tended to have comparable scores or slow slightly worse performance at post-test ($r = -.488$, $p < .001$), See Figure A3.2. No other main effects or interactions reached significance, *p's* > .2.

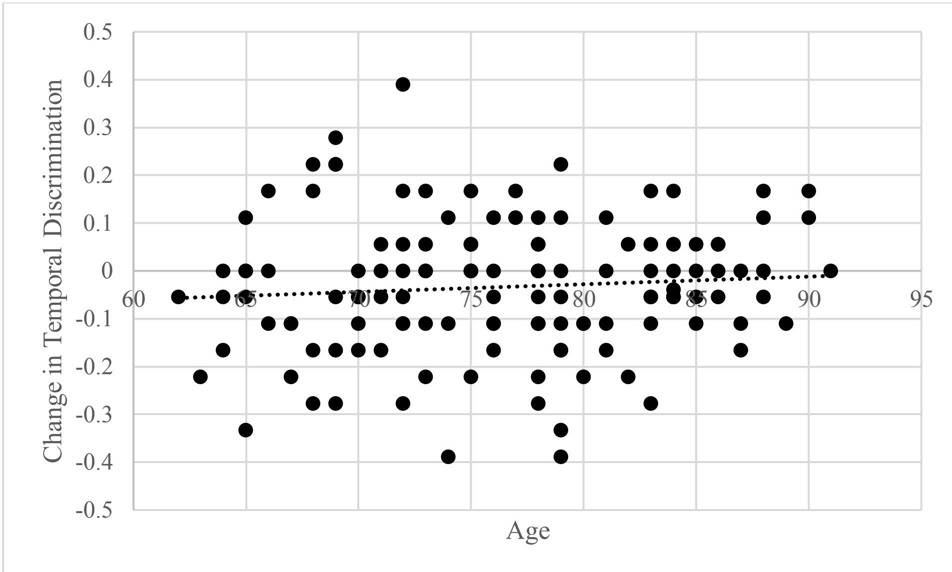

Figure A3.1. Younger children tended to show worse performance at post-test.

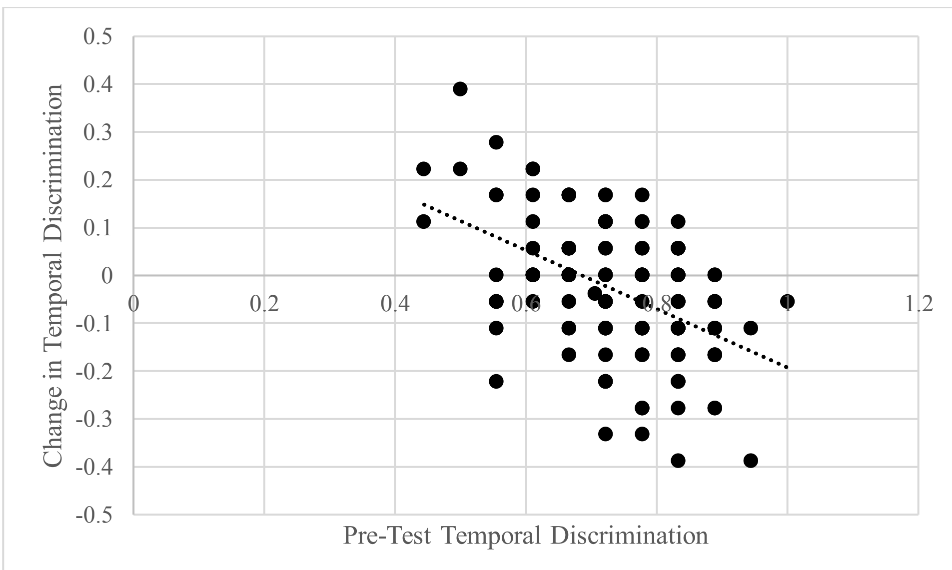

Figure A3.2. Children who performed poorly at pre-test tended to improve at post-test.

Sub-second timing. When including sub-second performance as a covariate, the 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA (including age and pre-test sub-second timing performance as covariates) on sub-second performance revealed a Time x Pre-test performance interaction, $F(1, 124) = 61.381$, $p < .001$, $\eta_p^2 = .331$. As with overall temporal discrimination performance, children with worse performance at pre-test showed greater gains over time than children who performed well at pre-test (r = .578, $p < .001$), See Figure A3.3. No other main effects or interactions reached significance, *p's* > .05.

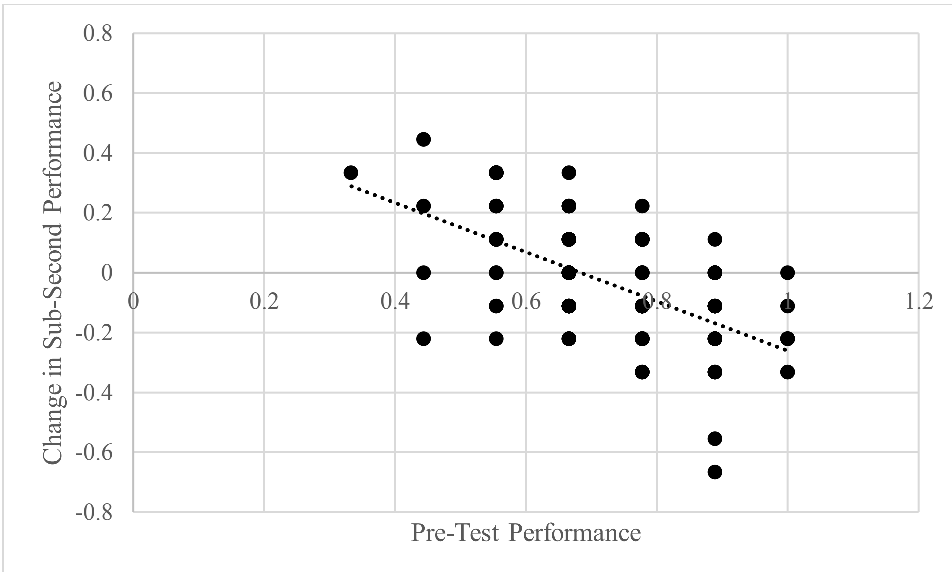

Figure A3.3. Children who performed poorly at pre-test showed greater gains.

Supra-second timing. When including supra-second performance at pre-test, the 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA (including age and pre-test supra-second performance as covariates) on supra-second performance showed a significant Time x Age interaction, $F(1, 127) = 7.973$, $p = .006$, $\eta_p^2 = .059$. Although the correlation between pre-test performance and age did not reach significance ($r = .094$, $p = .283$), a plot revealed that younger children tended to show no change in performance, whereas older children showed a slight improvement in performance, See Figure A3.4. There was also a significant Time x Pre-Test Performance interaction, $F(1, 124) = 62.575$, $p <.001$, $\eta_p^2 = .335$. Children with lower performance at pre-test showed greater improvements compared to children who started with better pre-test performance (r = - .594, p < .001), See Figure A3.5. No other main effects or interactions reached significance, p's > .3.

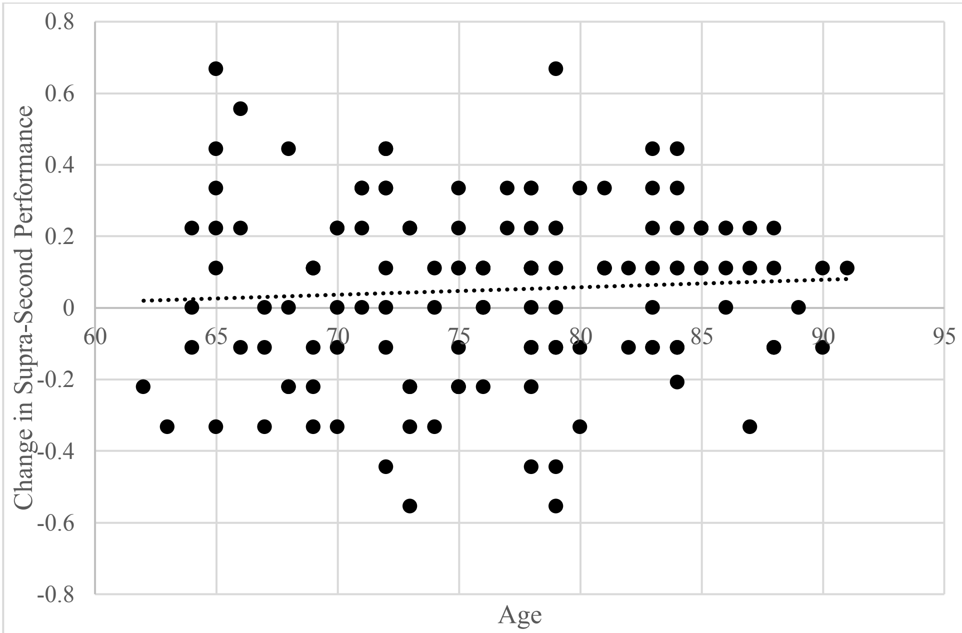

Figure A3.4. Younger children tended to have worse performance at post-test.

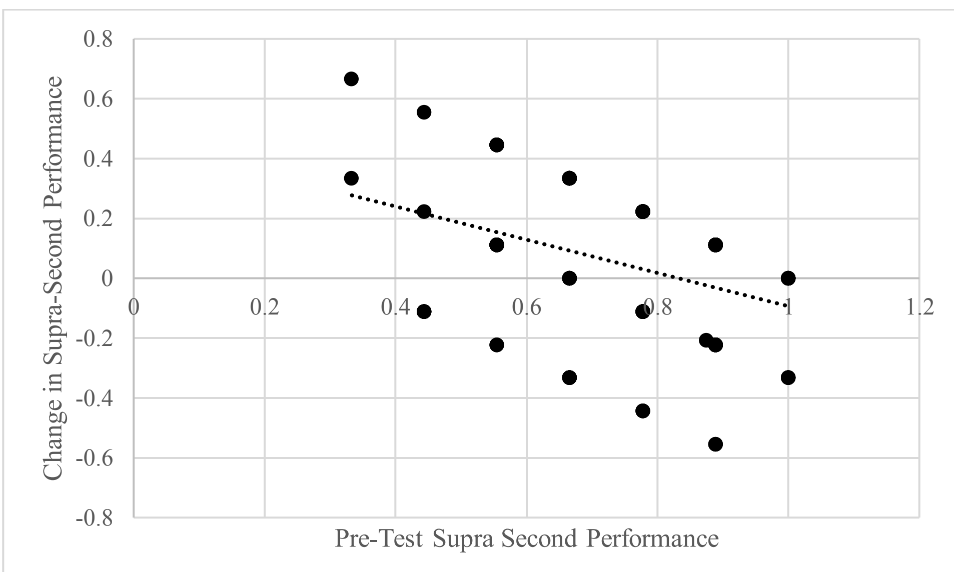

Figure A5. Children who performed lower at pre-test showed greater improvements in supra-second timing.

Temporal estimation (symbolic to nonsymbolic)

Average Estimate. The 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA (including age and pre-test estimate as covariates) on average estimate showed a main effect of Time, $F(1, 78) = 174.82$, $p < .001$, $\eta_p^2 = .691$, such that estimates were smaller at post-test ($M = 4.17$, $SE = .150$) compared to pre-test ($M = 4.91$, $SE = .000$). A main effect of Condition also emerged, $F(2, 78) = 32.769$, $p < .001$, $\eta_p^2 = .457$. Those in the Strategies condition ($M = 5.39$, $SE = .136$) had significantly larger estimates than those in the Symbols condition ($M = 4.37$, $SE = .120$) and marginally larger estimates than those in the Control condition ($M = 3.86$, $SE = .136$, $p < .001$). Those in the Symbols condition had larger estimates than those in the Control conditions ($p = .025$), See Figure A3.6. Lastly, there was also a significant Time x Pre-Test interaction, $F(1, 78) = 361.\ 846$, $p < .001$, $\eta_p^2 = .823$. Children who provided larger estimates at pre-test did not change their estimates as much as children who provided lower estimates ($r = -.864$, $p < .001$), See Figure A3.7. No other main effects or interactions reached significance, *p's* > .2.

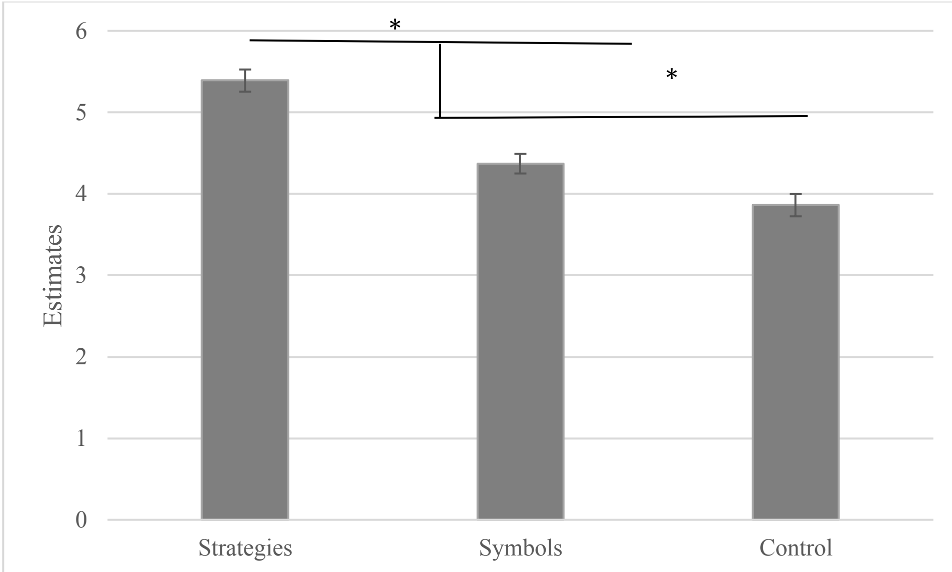

Figure A3.6. Those in the Strategies condition had the largest estimates, followed by those in the Symbols and then Control conditions.

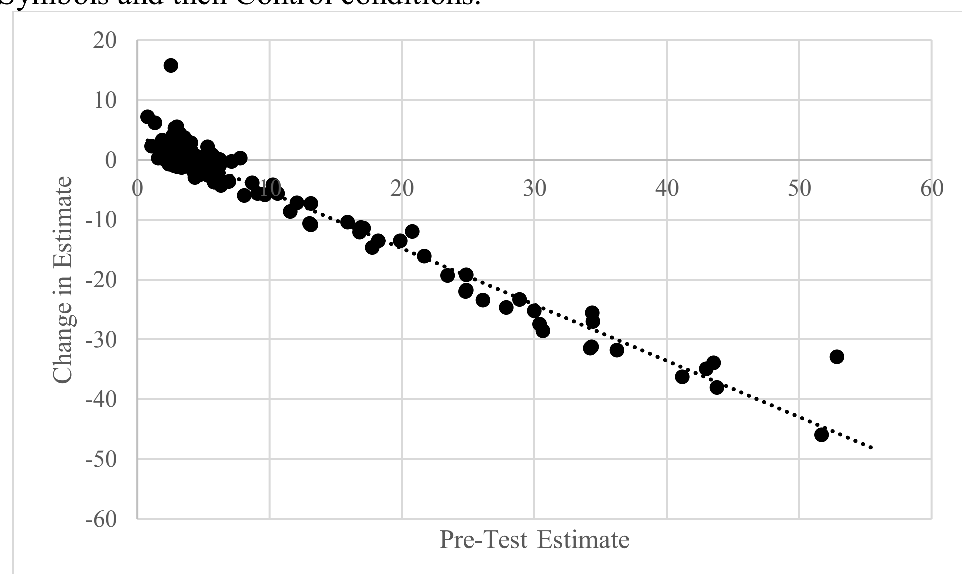

Figure A3.7. Children who provided larger estimates at pre-test were less likely to change their estimates significantly.

Slope. The 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA (including age and pre-test slope as covariates) on slope revealed a main effect of Condition, $F(2, 122) = 11.991$, $p < .001$, $\eta_p^2 = .164$, that was qualified by a Time x Condition interaction, $F(2, 122) = 11.991$, $p < .001$, $\eta_p^2 = .164$.

All conditions showed shallower slopes at post-test compared to pre-test (Strategies: Pre-Test: M = 1.138, SE = .001; Post-Test: M = .939, SE = .072, p = .007; Symbols: Pre-Test: M = 1.138, SE = .00; Post-test: M = .621, SE = .071, p <.001; Control: Pre-Test: M = 1.138, SE = .00; Post-Test: M = .444, SE = .072, p < .001). However, those in the Strategies condition had a steeper slope at post-test compared to those in the Symbols (p = .002) and Control conditions (p < .001). The slopes of those in the Symbols and Control conditions did not differ significantly (p = .085), See Figure A3.8. There was a Time x Pre-Test interaction, $F(1, 122) = 3052.990$, $p < .001$, $\eta_p^2 = .962$. A negative correlation between pre-test slope and change in slope over time (r = -.816, p < .001), and plotting the data showed that children with a negative slope at pre-test showed a positive change in slope whereas children with a positive slope at pre-test showed a negative change in slope, See Figure A3.9. No other main effects or interactions reached significance, *p's* > .4.

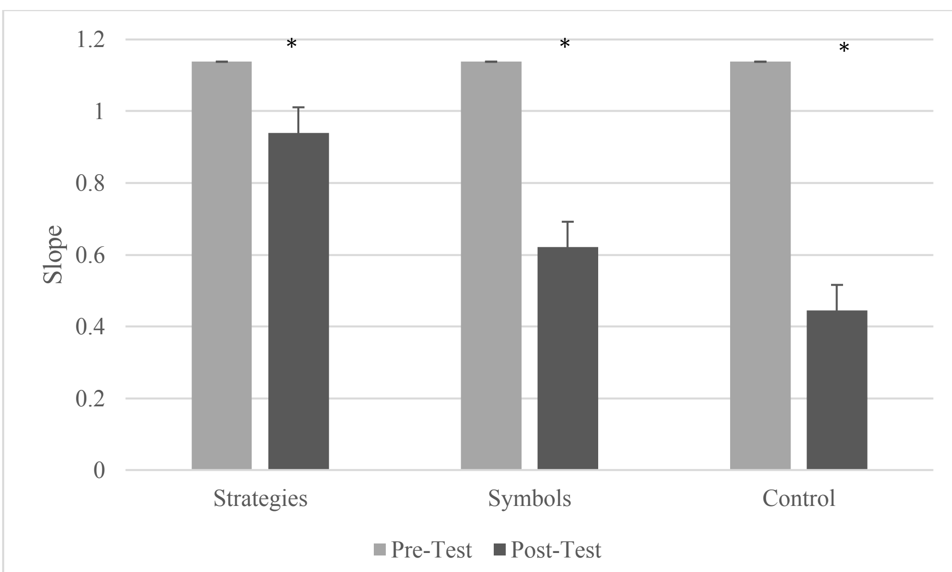

Figure A3.8. Those assigned to the Strategies condition had steeper slopes at pre-test, whereas those assigned to the Symbols condition had shallower slopes at pre-test.

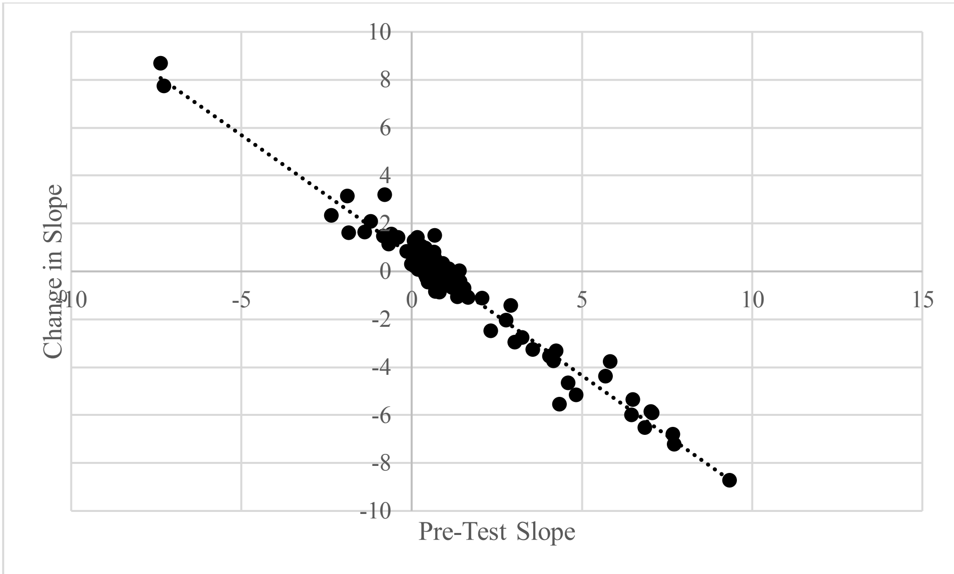

Figure A3.9. Children who had steeper slopes at pre-test showed less movement in their slopes at post-test.

Temporal Estimation (Nonsymbolic to Symbolic)

Error. The 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA (including age and pre-test error as covariates) on error showed a main effect of Condition, $F(2, 121) = 15.378$, $p < .001$, $\eta_p^2 = .203$, that was qualified by a significant Time x Condition interaction, $F(2, 121) = 15.378$, $p < .001$, $\eta_p^2 = .203$. Those in the Strategies Condition (Pre-Test: $M = .734$, $SE = .000$; Post-Test: $M = .326$, $SE = .083$; $p < .001$) produced less error at post-test, whereas those in the Control condition had greater error at post-test (Pre-Test: $M = .734$, $SE = .000$; Post-Test: $M = .943$, $SE = .083$; $p = .013$). Those in the Symbols condition produced comparable error at both time points (Pre-Test: $M = .734$, $SE = .000$; Post-Test: $M = .808$, $SE = .083$; $p = .374$), See Figure A3.10. Finally, there was a Time x Pre-Test interaction, $F(1, 121) = 32.433$, $p < .001$, $\eta_p^2 = .211$. A negative correlation between pre-test error and change in error ($r = -.445$, $p < .001$), and plotting the data indicated that those with greater error had pre-test had less error post-test, whereas those with little error at pre-test also had a minimal amount of error at post-test, See Figure A3.11. No other main effects or interactions reached significance, *p's* > .08.

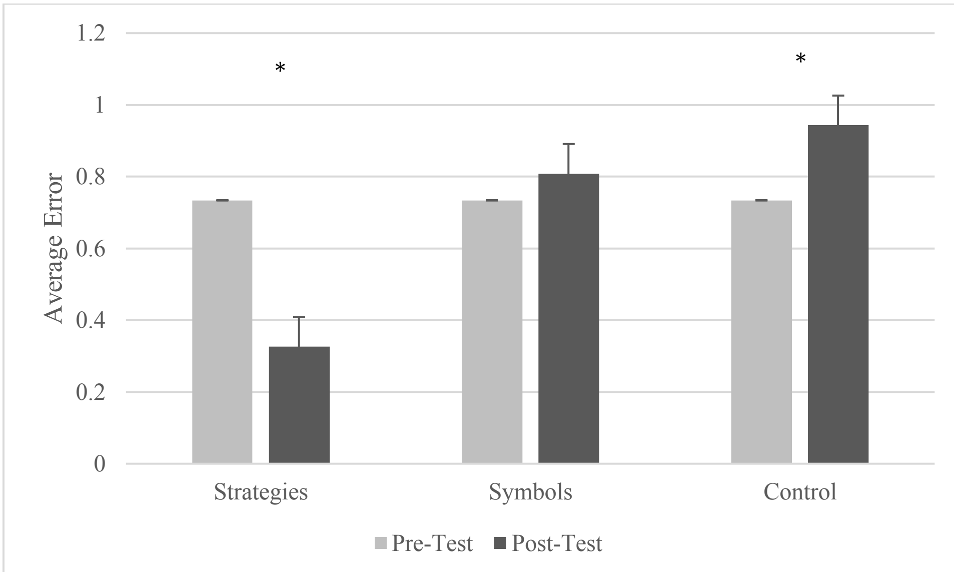

Figure A3.10. Children in the Strategies and Symbols condition had less error at post-test.

A3.11. Children with little error at pre-test also had a small amount of error at post-test.

Numerical Discrimination. The 2 (Time: Pre- versus Post-Test) x 3 (Condition: Strategies, Symbols, Control) mixed measures ANCOVA (including age and pre-test numerical discrimination performance as covariates) on numerical discrimination performance showed a main effect of Time, $F(1,123) = 24.572$, $p < .001$, $\eta_p^2 = .167$, such that performance was slightly worse at post-test ($M = 92.34$, $SE = .602$) compared to pre-test ($M = 92.82$, $SE = .000$). There was also a Time x Pre-Test Performance interaction, $F(1, 123) = 52.336$, $p < .001$, $\eta_p^2 = .298$. Again, a negative correlation between pre-test performance and change in performance over time ($r = -.536$, $p < .001$) and plotting this relation revealed that children with higher performance at pre-test showed no improvement (or worse performance) at post-test, whereas children with lower pre-test scores performed better at post-test, See A3.12. No other main effects or interactions reached significance, *p's* > .1.

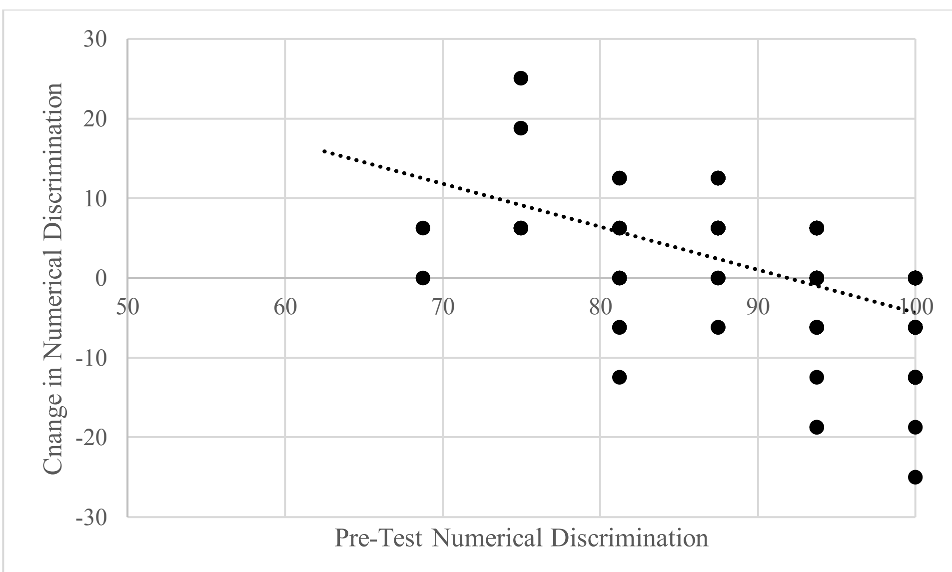

Figure A3.12. Children with lower pre-test performance showed greater improvements at post-test.

CHAPTER 4: MANIPULATING ADULTS' SYMBOLIC MAPPING OF TIME DOES NOT IMPACT TEMPORAL ACUITY

Abstract

Recent work suggests that children's understanding of temporal units of measurement is correlated with nonsymbolic temporal acuity (Hamamouche & Cordes, accepted). Whether this relation holds in adults - after temporal symbols have been mastered - and whether this correlation may be indicative of a causal relation, is unknown. In the current study, we explore whether temporal symbol understanding causally impacts temporal acuity in adulthood. Across two experiments, adults were given feedback to influence their estimates of temporal durations, and we explored whether this feedback influenced temporal discrimination abilities. Notably, in contrast to findings with children, results revealed no correlation between adults' nonsymbolic (temporal discrimination) and symbolic (temporal estimation) abilities. Although feedback successfully shifted adults symbolic and nonsymbolic timing estimates, changing one's mapping had no impact on subsequent temporal discrimination performance. Together these studies suggest that nonsymbolic and symbolic timing are less strongly related in adulthood, compared to childhood. Future work will be necessary for determining whether the relation between symbolic and nonsymbolic timing may be more malleable in childhood, when temporal symbols are being acquired, and also whether a bidirectional relation may exist in children who are still developing symbolic representations of time.

Over the course of development, children acquire several symbolic measurement systems that allow them to assign exact, discrete measurements to quantities they had previously only been able to track approximately. For example, preverbal infants successfully discriminate large changes in set sizes (e.g., Lipton & Spelke, 2003; Xu & Spelke, 2000; Xu & Arriaga, 2007), but it is only when children learn to verbally count that they can represent sets of objects exactly. Similarly, although infants notice two-fold changes in duration (e.g., Brannon, Suanda, & Libertus, 2007; vanMarle & Wynn, 2006), it is not until they learn the meaning behind the words "second" and "minute" that they understand how to track time precisely. Significant research dedicated to understanding the relation between nonsymbolic (i.e., approximate) and symbolic (i.e., exact) quantitative abilities has revealed a unique relation between these abilities within the domains of time and number (e.g., number: Bonny & Lourenco, 2013; Halberda, Mazzocco, & Feigenson, 2008; Libertus, Feigneson, & Halberda, 2011; for meta-analyses see: Chen & Li, 2014; Schneider et al., 2017; time: Hamamouche & Cordes, accepted) such that individuals with a better understanding of formal number symbols or temporal measurements are also better at discriminating nonsymbolic number/time. These correlational studies however, have been unable to address whether these relations are causal. That is, it is an open question whether having a more precise ability to track quantities nonsymbolically may lead to earlier/better learning of these symbols (scaffolding hypothesis, e.g., Starr, Libertus, & Brannon, 2013), and/or whether learning quantitative symbols may contribute to more precise nonsymbolic tracking abilities (refinement hypothesis, e.g., Mussolin, Nys, Content, & Leybaert, 2014). The current study aims to specifically address the refinement hypothesis within the domain of time by determining whether manipulating adults' symbolic mapping of time shape their temporal discrimination abilities.

A plethora of research has found a relation between an understanding of numerical symbols (i.e., math) and nonsymbolic numerical acuity (see Chen & Li, 2014; Schneider et al., 2017). This relation holds across different age ranges and measures of symbolic number understanding. For example, infants' approximate number abilities predict preschool-aged math performance (Starr, Libertus, & Brannon, 2013) and preschoolers' performance on a numerical discrimination task predicts school aged math (Mazzocco, Feigenson, & Halberda, 2011). Less work has investigated whether a similar relation holds in the domain of time. In one recent study, it was found that children's understanding of temporal symbols and their performance on a nonsymbolic temporal discrimination task was significantly correlated, even when controlling for age and numerical abilities (Hamamouche & Cordes, accepted). It is important to note, however, that the ages of the children in this study were specifically chosen in order to explore this relation at the point in development when children are learning about temporal symbols (e.g., 1st-3rd grades; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). Given the significant variability of symbolic knowledge in this age group, it is possible that the development of other cognitive factors, such as executive functioning, academic achievement, etc., may have contributed to this relation. Whether this relation holds in adulthood, when cognitive maturation stabilizes, remains an open question. Moreover, the study was correlational in nature, preventing us from making any causal claims. In the present study, we not only investigate whether this correlation holds in adulthood, but also whether a causal relation exists such that learning temporal units of measurement enhances nonsymbolic representations of time.

How might learning temporal units of measurement impact nonsymbolic representations of time? Prior to acquiring symbols, one has noisy representations of time. For example, one might perceive any duration between .5-2 seconds as being approximately one second. Acquiring an exact representation of that duration, however, may lessen the amount of noise in the underlying representation. Importantly, this exact measurement likely allows individuals to represent time symbolically even when performing nonymbolic temporal discrimination tasks. As such, it seems likely that learning temporal units of measurement would refine temporal acuity. Although this prediction has yet to be tested in the domain of time, the numerical cognition literature supports this possibility. For example, math abilities (i.e., symbolic representations of number) predict later nonsymbolic numerical performance, even when controlling for earlier nonsymbolic numerical abilities (Lyons, Bugden, Zheng, De Jesus, & Ansari, 2018; Matejko & Ansari, 2014; Mussolin et al., 2014; Suárez-Pellicioni & Booth, 2018). Moreover, one study found a sharp increase in preschoolers' numerical acuity following an acquisition of the cardinal principle – reflecting mastery of the verbal counting procedure (Shusterman et al., 2016). While these investigations hint at the possibility of a causal relation, no work has experimentally manipulated symbolic quantity abilities to determine whether our understanding of quantitative symbols may directly shape nonsymbolic acuity.

In the present study, adult participants completed a series of timing tasks to explore the relation between temporal symbols and nonsymbolic timing in adulthood. However, because adults have already acquired a temporal system of measurement we did not attempt to teach them new symbols, but instead provided feedback to shift their understanding of the durations represented by those symbols (i.e., to modify the mapping between symbolic and nonsymbolic time). In both experiments, participants were given a temporal discrimination task (nonsymbolic timing assessment) before and after a symbolic (Experiment 1) or nonsymbolic (Experiment 2) temporal estimation task in which they were provided with one of three types of feedback (overestimation, underestimation, or control) regarding their temporal estimates. Across both experiments, we investigated whether shifting either symbolic or nonsymbolic mappings had an impact on later temporal discrimination abilities.

How would modifying the mapping between symbolic and nonsymbolic time impact adults' subsequent nonsymbolic temporal acuity? We suggest two possibilities. First, providing feedback may shift the mapping between symbolic and nonsymbolic timing, resulting in a consistently higher (or lower) encoding of nonsymbolic temporal durations. That is, after receiving feedback that temporal estimates are too long (overestimation feedback), then a participant may lower their symbolic understanding of each duration by a constant amount. For example, they may map their understanding of 2 seconds onto a slightly shorter duration, thus encoding a mental magnitude of 1.5 sec as actually representing 2 secs. Following this adjustment, if the mapping between temporal magnitudes and symbols plays a role in nonsymbolic timing abilities, this participant may perceive temporal intervals in the discrimination task as being shorter, resulting in easier discriminations (e.g., if durations are encoded as consistently 0.5 sec shorter, then a comparison of 2 secs vs. 3 secs (.67 ratio) may be *perceived* as 1.5 sec vs. 2.5 sec (.6 ratio)).

Similarly, a participant receiving feedback indicating that they consistently underestimated the value (underestimation feedback) may find temporal discriminations to be more difficult following an adjustment of mapping durations to symbolic representations of longer durations. Given that adults are known to typically underestimate visually presented durations (e.g., Danckert & Allman, 2005; Lustig & Meck, 2011), a second possibility is that providing feedback to skew estimates closer to accurate estimates of duration (i.e., resulting in longer estimates) may result in more precise timing abilities. In this instance, we would expect to find temporal discrimination in the underestimation condition, and either unaffected, or worsened, temporal discrimination in the overestimation condition. In Experiment 1, we explore the relation between symbolic and nonsymbolic timing. Results reveal that shifting adults' symbolic estimates does not impact temporal acuity. In Experiment 2, we follow up on this finding by shifting adults' nonsymbolic mappings during a nonsymbolic estimation task; however, we again see no impact on a subsequent temporal acuity.

Experiment 1

In Experiment 1, participants completed a temporal discrimination task, and then a symbolic temporal estimation task. Halfway through the estimation task, participants were told they were either consistently overestimating, underestimating, or fairly accurately estimating the requested durations, and were asked to adjust their estimates accordingly. Following the temporal estimation task, participants again completed the temporal discrimination task, to explore how feedback on the estimation task influenced nonsymbolic temporal abilities.

Methods

Participants

Undergraduates (N = 196) participated in this study for course credit or cash compensation. A surprising number of participants provided responses in the temporal estimation task that were not within the realm of reasonable responses (estimating durations between 1181-2300ms to be less than 10 ms at either pre- or post-test. These participants were excluded from the analyses (n = 52) on the basis that they did not know the duration of a millisecond. An additional 13 participants completed the study, but were excluded due to program error. Thus, the final sample included 131 adults ($n_{\text{Overestimation}}$ = 43; $n_{\text{Underestimation}}$ = 44; n_{Control} = 44, M_{age} = 18.97 years, 90 females).

Procedure

Testing took place in a quiet room in the laboratory. Adults completed the following three tasks in a single testing session: 1) temporal discrimination (nonsymbolic timing), 2) temporal estimation with overestimation, underestimation, or control feedback (symbolic timing), 3) temporal discrimination (nonsymbolic timing).

(1) Temporal discrimination

During the temporal discrimination task, participants saw two characters (a chicken and a cow) presented on the computer screen. On each trial, each character was shown playing a musical instrument (chicken always went first) in which a sound was played (horn sound for chicken, train sound for cow) and a box appeared around the animal for the duration of the instrument's sound. Participants were instructed to press keys on the keyboard to indicate which animal played an instrument longer. Given that adults are better at discriminating supra-second durations compared to sub-second durations (Hamamouche & Cordes, 2019), half of the comparisons were in the sub-second range (< 1000ms), and half of the comparisons were in the supra-second range (> 1000ms). The durations were randomly intermixed throughout the session, and the chicken played the instrument longer on approximately half of the trials. Each comparison differed by one of the following ratios: 1.14, 1.2, 1.5, and 2. Adults completed 32 trials: 2 duration ranges x 4 ratios x 4 times each.

(2) Temporal estimation with feedback

After completing the temporal discrimination task, participants began the temporal estimation task during which an oval was displayed on the computer screen for one of five specified durations (1181, 1461, 1741, 2020, and 2300 ms). After the oval disappeared, participants were asked to type the duration of the oval (in milliseconds) into a textbox. After completing 30 baseline trials (5 durations, 6 times each), a message appeared notifying the participants of their performance. This feedback was based on the condition (overestimation, underestimation, or control) to which the participant was randomly assigned, and thus all feedback was pre-determined and did not necessarily reflect the participant's true performance. Participants assigned to the overestimation condition saw the message, "You're doing a good job, but, on average, you are OVERESTIMATING by 391 milliseconds. The goal is to be as accurate as possible. Please adjust your estimates so that they are slightly SHORTER in order to be more accurate." Participants assigned to the underestimation condition were told, "You're doing a good job, but, on average, you are UNDERESTIMATING by 391 milliseconds. The goal is to be as accurate as possible. Please adjust your estimates so that they are slightly LONGER in order to be more accurate." In a third control condition, participants saw the message, "You're doing a good job. Please continue." The feedback was specific to the participant's assigned condition and did not reflect the participant's true performance. Following the feedback, the participants completed another 30 trials of the temporal estimation task.

After completing the temporal estimation task, participants completed the temporal discrimination task again. At the end of the study, participants were debriefed (including an explanation of the deceptive feedback) and received course credit or cash compensation for their time.

Data Coding and Analyses

Temporal discrimination

The dependent measure for the temporal discrimination task (overall, sub-second, and supra-second) was proportion correct.

Temporal estimation

Average estimates were calculated before and after feedback and served as the dependent measure for the temporal estimation task. To look for outliers, we first checked for performance three standard deviations above or below the mean. One participant had an average estimate greater than three standard deviations above the mean, and was thus excluded from analyses involving estimation. As a secondary measure of temporal estimation performance, we calculated the slope between the durations presented and the estimates participants provided. The slope allowed us to determine whether participants increased their estimates with an increasing target duration. One additional participant had a slope greater than three standard deviations above the mean at pre-test and was thus excluded from analyses involving estimation performance.

Results

Table 4.1 shows the average performance for each task for each condition separately and Table 4.2 shows correlations between pre- and post-test performance for each condition separately.

	Pre-Test			**Post-Test**		
	Discrimination	**Estimation**		**Discrimination**	**Estimation**	
	Proportion Correct	**Estimate**	**Slope**	**Proportion Correct**	**Estimate**	**Slope**
Overestimation	.80 (.01)	1126.87 (150.30)	.67 (.09)	.82 (.01)	875.07 (136.02)	.56 (.09)
Underestimation	.81 (.01)	779.74 (141.99)	.43 (.07)	.84 (.01)	1089.27 (141.78)	.47 (.07)
Control	.78 (.01)	595.19 (118.70)	.37 (.08)	.80 (.01)	552.42 (117.99)	.29 (.06)

Table 4.1. Average performance (SE) on each task.

	Discrimination	**Estimation**	
	Proportion Correct	**Estimate**	**Slope**
Overestimation	.167	.954**	.902**
Underestimation	$.286^{+}$	.964**	.904**
Control	.078	.959**	.897**

Table 4.2. Correlations between pre- and post-test performance
$^{**}p < .001$, $^{*}p < .01$, $^{+}p < .1$

Effects of Feedback on Temporal Estimation

Temporal estimation (average estimate). Overall, participants underestimated the average estimate at pre-test (M = 829.43, SE = 80.97) compared to the actual average (M = 1740.6), $t(129) = -11.254$, $p < .001$. Although participants were randomly assigned to a feedback condition, there was a significant difference in the average estimate at pre-test, $F(2, 127) = 3.820$, $p = .024$. Follow-up independent samples t-tests revealed that participants in the overestimation condition provided significantly larger estimates at pre-test (M = 1126.87, SE = 150.30) compared to those in the control condition (M = 595.19, SE = 118.70), $t(84) = 2.79$, $p = .007$. There was no significant difference between the underestimation (M = 797.74, SE = 141.99) and overestimation conditions ($t(84) = 1.68$, $p = .097$), nor between the underestimation and control conditions ($t(86) = .997$, $p = .32$).

To ensure that adults shifted performance on the temporal estimation task based on the feedback received, we conducted a 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANOVA on the average temporal estimate. There was a marginal effect of Condition, $F(2, 127) = 3.01$, $p = .053, \eta_p^2 = .045$, which was qualified by a significant Time x Condition interaction, $F(2, 127) = 50.27$, $p < .001, \eta_p^2 = .442$. To better understand the interaction, we conducted separate paired samples t-tests on the average estimate at pre- and post-test for each condition separately. As instructed, participants assigned to the overestimation condition significantly shortened their estimates at post-test (Pre-Test: M = 1126.87, SE = 150.30; Post-Test: M = 885.82, SE = 138.86; $t(41) = 5.304$, $p < .001$) and participants assigned to the underestimation condition provided significantly longer estimates after receiving feedback (Pre-Test: M = 779.74, SE =141.99; Post-Test: M = 1089.27, SE =141.78; $t(43)$

= -8.11, p < .001). As expected, participants assigned to the control condition did not significantly change their estimates after feedback (Pre-Test: M = 595.19, SE =118.70; Post-Test: M = 552.42, SE =117.99; t(43) = 1.262, p = .214). The main effect of Time did not reach significance, F(1, 127) = .143, p = .706, η_p^2 = = .001. See Figure 4.1.

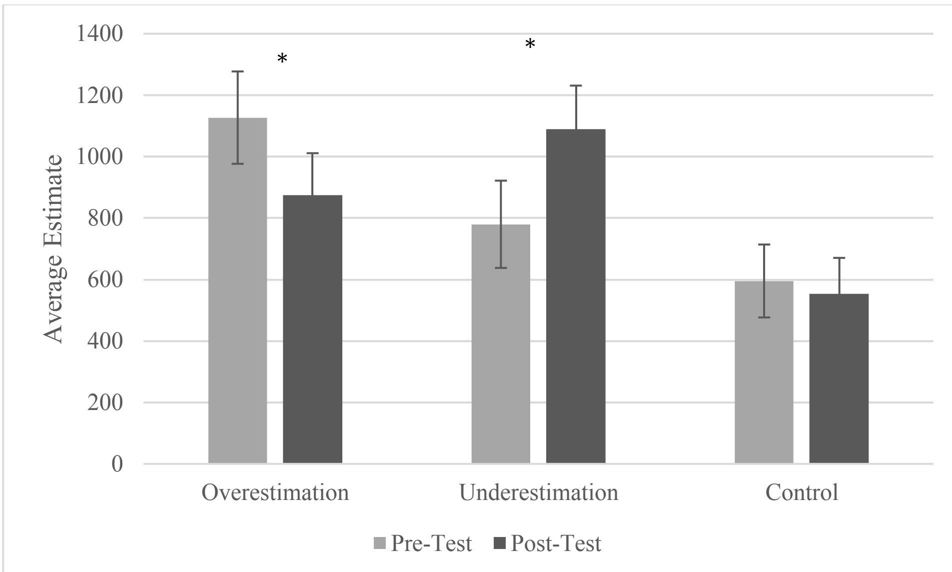

Figure 4.1. Average estimates on the temporal estimation task before and after feedback.

Temporal estimation (slope). Participants' slopes were significantly greater than 0 at pre-test (M = .49, SE = .05; t(128) = 10.087, p < .001), indicating a positive slope between participants responses and the oval's duration. Again, inexplicably, the one way ANOVA revealed significant differences in participants' slopes at pre-test, F(2, 126) = 3.831, p = .024. Those assigned to the overestimation condition (M = .67, SE = .09) initially had greater slopes than those assigned to the underestimation condition (M = .43, SE = .07, t(83) = 2.062, p = .042) and the control condition (M = .37, SE = .08, t(84) = 2.513, p = .014). There was no significant difference between the underestimation and control conditions, p = .572.

The 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANOVA on slope revealed a main effect of Time, $F(1, 126) = 5.559$, $p = .020$, $\eta_p^2 = .042$, such that the slope was significantly steeper at pre-test ($M = .488$, $SE = .047$) compared to post-test ($M = .439$, $SE = .043$). There was also a main effect of Condition, $F(2, 126) = 3.657$, $p = .029$, $\eta_p^2 = .055$. Those in the overestimation condition ($M = .61$, $SE = .09$) had significantly steeper slopes than those in the control condition ($M = .3304$, $SE = .07$, $t(85) = 2.471$, $p = .015$). Slopes in the underestimation condition fell between the two (p's >.18). Importantly, the Time x Condition interaction was significant, $F(2, 126) = 3.631$, $p = .029$, $\eta_p^2 = .054$. Slopes were shallower at post-test compared to pre-test for those assigned to the overestimation (Pre-Test: $M = .67$, $SE = .09$; Post-Test: $M = .57$, $SE = .09$; $t(41) = 2.559$, $p = .014$) and control (Pre-Test: $M = .37$, $SE = .08$; Post-Test: $M = .29$, $SE = .06$; $t(43) = 2.026$, $p = .049$) conditions. Slopes were comparable at pre- ($M = .43$, $SE = .07$) and post-test ($M = .46$, $SE = .07$; $t(42) = -.896$, $p = .375$) for those assigned to the underestimation condition, indicative of a constant change in estimates. See Figure 4.2.

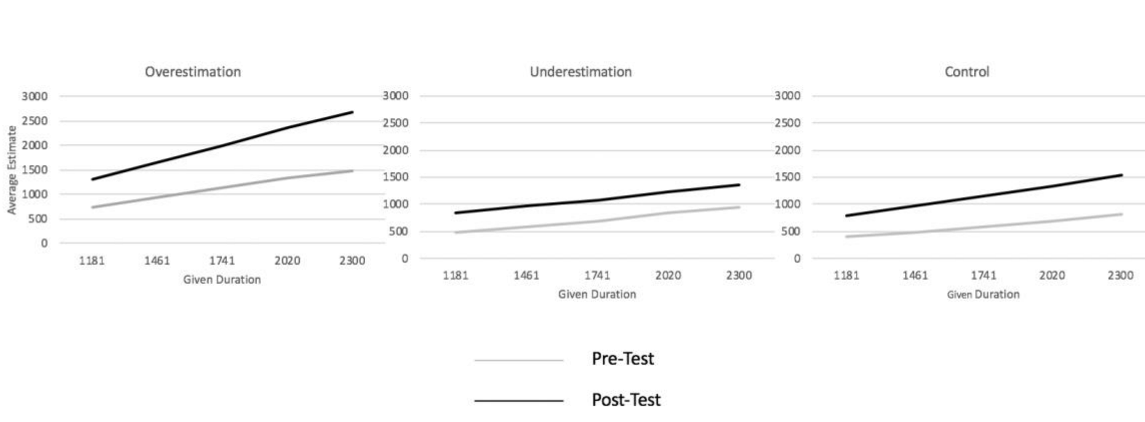

Figure 4.2. Changes in participants' slopes on the estimation task before and after feedback.

Effects of Feedback on Temporal Estimation

Temporal discrimination. Prior to feedback, there were no condition differences on the temporal discrimination task, $F(2, 128) = 1.545$, $p = .217$, indicating that any condition differences after feedback were not due to pre-existing differences in the conditions. Temporal discrimination performance was above chance at pre- ($M = .79$, $SE = .006$; $t(130) = 48.64$, $p < .001$) and post-test ($M = .82$, $SE = .007$; $t(130) = 47.833$, $p < .001$). Neither average estimate ($r = .118$, $p = .180$), nor slope ($r = .113$, $p = .201$) was correlated with temporal discrimination performance at pre-test.

Previous work has shown that adults are more accurate at timing durations longer than a second (i.e., supra-seconds) compared to durations shorter than a second (i.e., sub-seconds; Hamamouche & Cordes, 2019). Moreover, temporal discriminations become easier as the ratio between the two durations increases (i.e., ratio effects, Brannon, Suanda, & Libertus, 2007; Odic, Lisboa, Eisinger, Olivera, Maiche, & Halberda, 2016; Odic, 2018). To determine whether similar patterns emerged in our data, we conducted a 2 (Duration Range: Sub- versus Supra) x 4 Ratio (1.14, 1.2, 1.5, 2) repeated measures ANOVA on the temporal discrimination data at pre-test. Replicating previous work, there was a main effect of Duration Range, $F(1, 130) = 28.285$, $p < .001$, $\eta_p^2 = .179$, such that performance was significantly better in the supra-second range ($M = .83$, $SE = .01$) compared to the sub-second range ($M = .75$, $SE = .01$). There was a main effect of Ratio, $F(3, 390) = 134.63$, $p < .001$, $\eta_p^2 = .509$, such that performance improved as the ratio became larger (all *p's* <.01). There was also a Duration Range x Ratio interaction, $F(3, 390) = 7.856$, $p < .001$, $\eta_p^2 = .057$. Follow-up paired samples t-tests between sub- and supra-second timing for each ratio separately revealed that performance was significantly better in the supra-second compared to the sub-second range for the three most difficult ratios (Means (SE): Ratio 1.14: $M_{Sub-Second}$ = .61 (.02) , $M_{Supra-Second}$ = .71 (.02), $t(130) = -3.662$, $p < .001$; Ratio 1.2: $M_{Sub-Second}$ = .65 (.02), $M_{Supra-Second}$ = .77 (.02), $t(130) = -4.695$, $p < .001$; Ratio 1.5: $M_{Sub-Second}$ = .82 (.02), $M_{Supra-Second}$ = .91 (.01), $t(130) = -4.469$, $p < .001$; however, for the easiest ratio, performance was comparable in both ranges ($M_{Sub-Second}$ = .93 (.01) , $M_{Supra-Second}$ = .92 (.01), $t(130) = .669$, $p = .504$). See Figure 4.3.

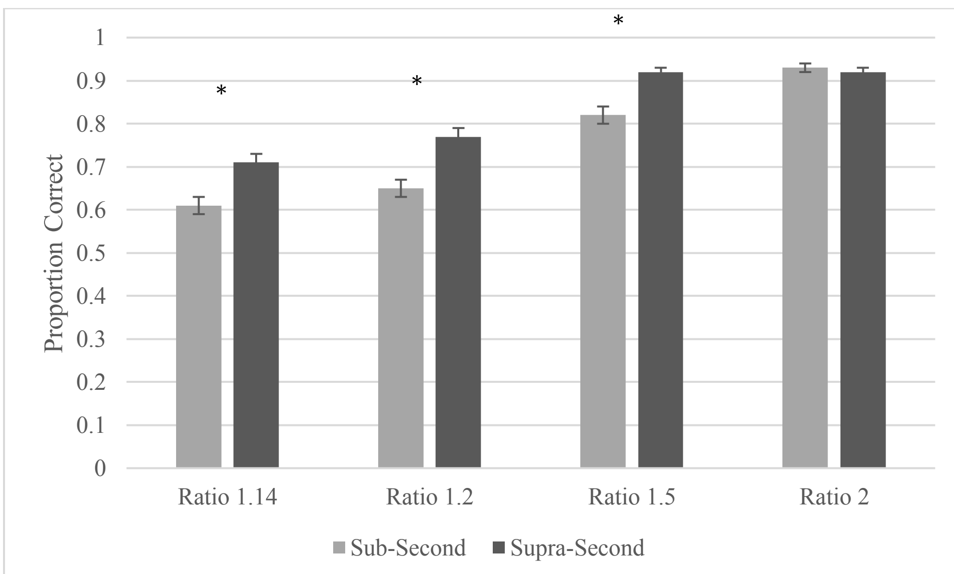

Figure 4.3. Performance was significantly better in the supra-second range compared to the sub-second range for all ratios, except for the easiest ratio.

To determine whether shifting adults' perceptions of time on the temporal estimation task impacted overall performance on temporal discrimination, we conducted a 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANOVA on temporal discrimination performance. There was a main effect of Time, $F(1, 128) = 7.224$, $p = .008$, $\eta_p^2 = \ = .053$, revealing that performance was significantly better at post-test (M = .82, SE = .007) compared to pre-test (M = .80, SE = .006). There was also a main effect of Condition, $F(2, 128) = 4.124$, $p = .018$, $\eta_p^2 = = .061$. Follow-up independent samples t-tests revealed that performance in the underestimation condition (M = .82, SE = .008) was significantly better than performance in the control condition (M = .79, SE = .008, $t(86) = 2.903$, $p = .005$). Performance did not differ between the overestimation (M = .81, SE = .008) and underestimation conditions ($t(85) = -1.324$, p = .189), nor between the overestimation and control conditions ($t(85) = 1.513$, $p = .134$). The Time x Condition interaction did not reach significance, $p > .5$. See Figure 4.4[6].

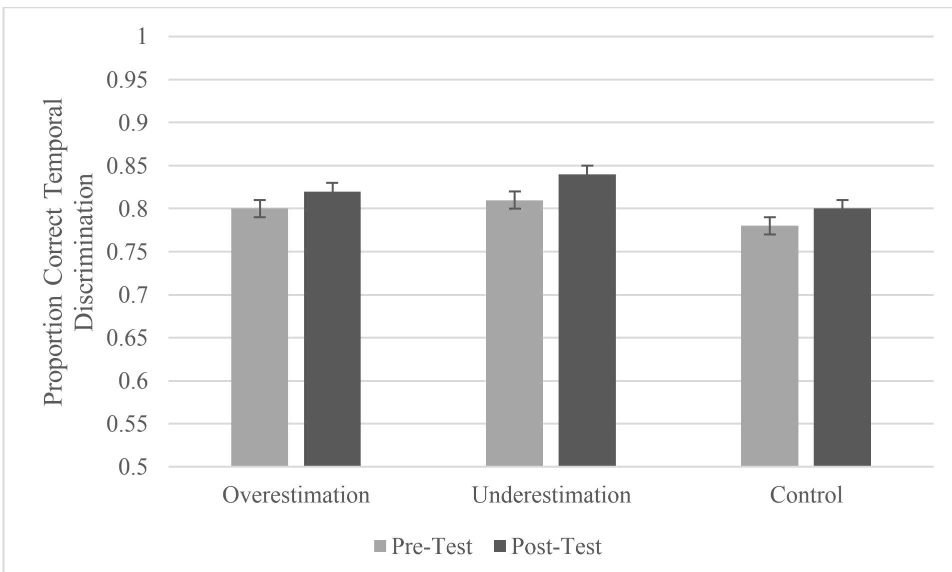

Figure 4.4. Across all three conditions, temporal discrimination performance improved after receiving feedback.

Finally, we conducted correlations between change in average estimate and post-test temporal discrimination, change in average slope and post-test temporal discrimination, change in average estimate and change in temporal discrimination, and change in slope and change in temporal discrimination. In order to test whether providing

[6] We also explored whether the feedback differentially affected sub- or supra-second timing. Neither ANOVA revealed a significant Time x Condition interaction, p's > .2

feedback that resulted in more precise symbolic estimates (i.e., underestimation feedback), we conducted the correlations for the underestimation and overestimation conditions separately; however, none of the correlations reached significance, *p's* > .1.

Discussion

The first goal of this study was to determine whether a correlation exists between symbolic and nonsymbolic timing in adulthood. Our results mirror prior work in the numerical cognition literature indicating that although this relation exists in children, it did not hold when using an adult sample (see Inglis, Attridge, Batchelor, & Gilmore, 2011 for a similar pattern of results with number). Why would a correlation exist in children, but not adults? The acquisition of strategies over the course of development may explain this difference. While children are in the process of acquiring and mastering symbolic representations of quantity, children may have a greater reliance upon nonsymbolic representations. However, as individuals become more experienced, they may begin to use more advanced strategies that rely less upon nonsymbolic representations and more upon explicit timing strategies, such as counting.

Our second aim was to study if shifting adults' symbolic mapping subsequently shaped their nonsymbolic representations of time. Notably, our feedback manipulation worked: participants told that they were overestimating provided significantly smaller estimates following feedback, and those told they were underestimating provided significantly longer estimates following feedback. Moreover, slope analyses revealed that those in the overestimation and control conditions adjusted estimates proportionally, whereas those in the underestimation condition adjusted their estimates in a constant fashion. Given that only constant changes in mapping should impact subsequent temporal discrimination performance, we would have expected a change in temporal discrimination performance in the underestimation condition only. Despite this, the effect of feedback did not extend to performance on the temporal discrimination task regardless of condition.

There are many explanations as to why the temporal discrimination performance was unaffected by feedback. First, given that we did not find a correlation between adults' nonsymbolic and symbolic timing, it seems unlikely that manipulating the mapping between symbolic and nonsymbolic representations would subsequently affect nonsymbolic representations. An alternative possibility is that our feedback manipulation was not strong enough. That is, our feedback manipulation may not have been convincing enough to impact the actual mapping between symbolic and nonsymbolic representations of time. Whether greater transfer between tasks would occur following more extensive feedback and practice remains an open question. Additionally, it is possible that feedback on a symbolic task may not readily transfer to a second, nonsymbolic task. In other words, it may be unlikely that shaping the mapping of symbols transfers to nonsymbolic performance as the symbolic and nonsymbolic task have distinct task demands. If this were the case, one might expect feedback on a nonsymbolic task to more readily transfer to another nonsymbolic task. Thus, in Experiment 2, we provide feedback on a nonsymbolic temporal reproduction task to determine whether manipulating one's nonsymbolic mapping impacts performance on a separate nonsymbolic task.

Experiment 2 also allowed us to investigate how calibrating one's internal clock might impact timing abilities. Previous work has shown that a dose of methamphetamine speeds up rats' internal clocks, resulting in temporal overestimation (Meck & Church, 1983). Similarly, timing of filled intervals (i.e., continuously presented stimuli), which are thought to be more arousing than timing empty intervals (i.e., timing an interval that is denoted by two individual stimuli), increases the rate of one's internal clock and subsequently leads to temporal overestimation (e.g., Thomas & Brown, 1974; Wearden, Norton, Martin, Montford-Bebb, 2007). By providing feedback on a nonsymbolic task, participants will be able calibrate their own internal clock allowing us to test whether feedback on a basic estimation task influences subsequent discrimination abilities.

Experiment 2

Experiment 2 was identical to Experiment 1 except that participants received feedback on a nonsymbolic temporal *reproduction* task, instead of a symbolic temporal estimation task.

Methods

Participants

Undergraduates (N = 126, 42 per condition, M_{age} = 19.39 years, 102 females) participated in this study for course credit or cash compensation. Required sample size was computed using a priori calculations in G*Power.

An additional four individuals participated in the study, but were excluded for program error (n = 2) or an incomplete data set (n = 2).

Procedure

Adults completed the following in a single testing session: 1) temporal discrimination, 2) temporal reproduction with feedback (overestimation, underestimation, or control feedback), 3) temporal discrimination. The procedure for Experiment 2 was identical to Experiment 1 except for the following: Instead of a temporal estimation task, participants completed a temporal reproduction task in which they saw an oval displayed on the computer screen (the same durations as Experiment 1: 1181, 1461, 1741, 2020, and 2300 ms) and were asked to hold down the space bar for the same amount of time that the oval had appeared. Identical to Experiment 1, participants received feedback on their performance after completing 30 baseline trials (5 durations, 6 times each), and then completed an additional 30 trials of the temporal reproduction task.

Data Coding

Temporal discrimination

The dependent measure for the temporal discrimination task (overall, sup-second, and supra-second) was proportion correct.

Temporal reproduction

We again used average estimate and slope as the dependent variables for the temporal reproduction task. First, we calculated average estimate before and after feedback as the dependent measure for the temporal reproduction task. Participants with performance three standard deviations above (Pre-Test: $n = 1$) or below (Post-Test: $n = 1$) the mean were excluded from analyses involving the average estimate. We next calculated slope. Participants with slopes three standard deviations above/below the mean were excluded from slope analyses (Pre-Test: $n = 2$, Post-Test: $n = 1$).

Results

Table 4.3 shows the average performance for each task and Table 4.4 shows correlations between pre- and post-test performance for each condition separately.

	Pre-Test			**Post-Test**		
	Discrimination	**Reproduction**		**Discrimination**	**Reproduction**	
	Proportion Correct	**Estimate**	**Slope**	**Proportion Correct**	**Estimate**	**Slope**
Overestimation	.81 (.01)	1339.47 (30.23)	.77 (.03)	.81 (.01)	1181.73 (37.52)	.63 (.03)
Underestimation	.80 (.01)	1360.10 (31.33)	.80 (.03)	.84 (.01)	1516.84 (42.57)	.78 (.03)
Control	.81 (.01)	1368.94 (34.07)	.78 (.03)	.82 (.01)	1392.84 (35.05)	.71 (.04)

Table 4.3. Average performance (SE) on each task.

	Discrimination	**Reproduction**	
	Proportion Correct	**Estimate**	**Slope**
Overestimation	.422*	.720**	.360^{+}
Underestimation	.445*	.547**	.380*
Control	.381*	.728**	.374*

Table 4.4 Correlations between pre- and post-test performance in each condition.
**$p < .001$, *$p < .01$, $^{+}p < .1$

Temporal reproduction (average estimate). Again, participants' average estimate during the reproduction task (M = 1356.30, SE = 18.32) was significantly less than the true average of the durations presented (M = 1740.6; $t(124) = -20.970$, $p < .001$). Performance was comparable across conditions on the temporal reproduction task prior to feedback, indicating there were no initial group differences, $F(2, 122) = .223$, $p = .801$.

Next, we determined whether participants changed their reproductions to match the feedback received during the temporal reproduction task. To do so, we conducted a 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANOVA on the average temporal reproduction. There was a main effect of Condition, $F(2, 121) = 98.653$ $p < .001$, η_p^2 =.125, which was qualified by a Time x Condition interaction, $F(2, 121) = 34.925$, $p < .001$, η_p^2 = .366. Follow-up paired samples t-tests revealed that as predicted, those assigned to the overestimation condition significantly lowered their estimates at post-test (Pre-Test: $M = 1339.47$, $SE = 31.98$; Post-Test: $M = 1181.73$, $SE = 37.40$; $t(40) = 6.009$, $p < .001$) whereas those assigned to the underestimation condition significantly raised their estimates at post-test (Pre-Test: $M = 1352.03$, $SE = 31.98$; Post-Test: $M = 1535.60$, $SE = 37.40$; $t(40) = -5.315$, $p < .001$). Those in the control condition had comparable estimates at both time points (Pre-Test: $M = 1368.94$, $SE = 31.60$; Post-Test: $M = 1392.84$, $SE = 36.95$; $t(41) = -.965$, $p = .340$), see Figure 4.5. The main effect of time did not reach significance, $p > .3$.

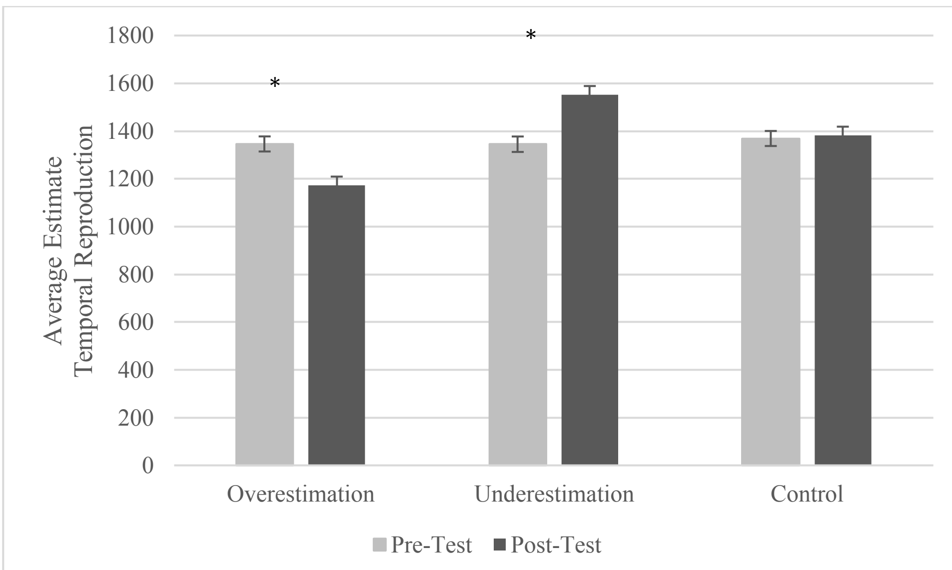

Figure 4.5. Average estimates on the temporal reproduction task before and after feedback.

Temporal reproduction (slope). Again, slopes were significantly greater than 0 at pre-test ($M = .78$, $SE = .02$), $t(123) = 47.22$, $p < .001$, indicating that participants' provided larger estimates as the target duration increased. There were no significant differences in reproduction slopes between conditions at pre-test, $F(2, 121) = .162$, $p = .851$.

We next conducted a 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANOVA on participants' slopes. There was again a main effect of Time, $F(1, 120) = 11.807$, $p < .001$, $\eta_p^2 = .090$, such that slopes were steeper at pre-test ($M = .785$, $SE = .017$) compared to post-test ($M = .713$, $SE = .021$). There was a significant Time x Condition interaction, $F(2, 120) = 3.435$, $p = .035$, $\eta_p^2 = .054$. Follow-up paired samples t-test revealed that those in the overestimation condition adjusted their estimates proportionally such that they had significantly steeper slopes at pre-test ($M = .78$, $SE = .03$) compared to post-test ($M = .63$, $SE = .03$, $t(39) = 3.912$, $p < .001$). However, those in the underestimation condition (Pre-Test: $M = .80$, $SE = .03$; Post-Test: $M = .78$, $SE = .03$; $t(42) = .509$, $p = .613$) and control condition (Pre-Test: $M = .78$, $SE = .03$; Post-Test: $M = .73$, $SE = .04$; $t(39) = 1.317$, $p = .195$) had comparable slopes at pre- and post-test, suggestive of a constant shift in temporal estimates. See Figure 4.6.

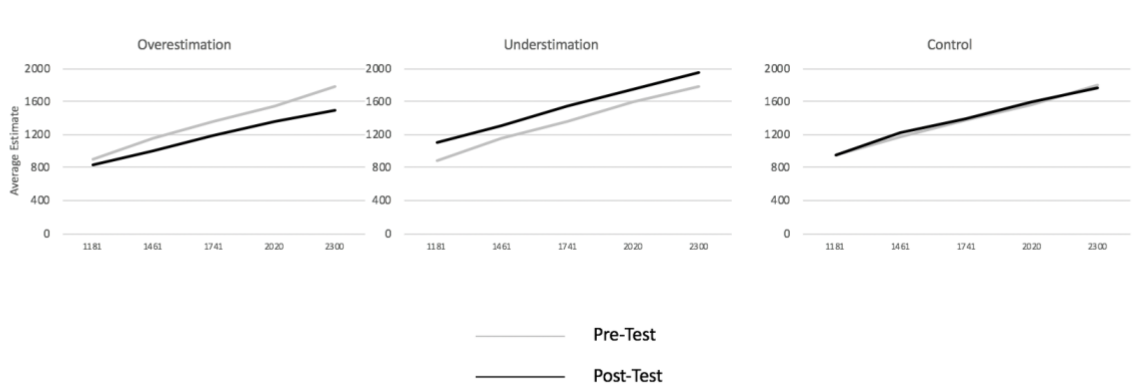

Figure 4.6. Slopes decreased after feedback in the overestimation condition only.

Temporal discrimination. There were no condition differences at pre-test on the temporal discrimination task, $F(2, 123) = .385$, $p = .681$. Performance was above chance on the temporal discrimination task at pre- and post-test (Pre-Test: $M = .81$, $SE = .007$; $t(125) = 44.316$, $p < .001$; Post-Test: $M = .82$, $SE = .007$; $t(125) = 43.645$, $p < .001$). Although there was no correlation between temporal discrimination performance and average reproduction during the temporal reproduction task, $r = .048$, $p = .598$, we replicated previous work showing a significant correlation between temporal discrimination and reproduction slope at pre-test, $r = .211$, $p = .019$ (Cicchini, Arrighi, Cecchetti, Giusti, & Burr, 2012).

We again tested for differences in sub- and supra-second timing and for ratio effects by conducting a 2 (Duration Range: Sub- versus Supra) x 4 Ratio (1.14, 1.2, 1.5, 2) repeated measures ANOVA on performance on the temporal discrimination task at pre-test. Replicating previous work (Hamamouche & Cordes, 2019) and Experiment 1, there was a main effect of Duration Range (Sub- versus Supra-), $F(1, 125) = 19.63$, $p < .001$, $\eta_p^2 = .136$) such that performance was significantly better in the supra-second range ($M = .83$, $SE = .009$) compared to the sub-second range ($M = .78$, $SE = .009$). There was also a main effect of Ratio, $F(3, 375) = 51.113$, $p < .001$, $\eta_p^2 = .290$, such that performance improved as the ratio became larger (all *p's* <.001, except between the two easiest ratios, $p = .756$). The Range x Ratio interaction also reached significance, $F(3, 375) = 15.143$, $p < .001$. $\eta_p^2 = .108$. Follow-up independent samples t-tests revealed that supra-second ($M = .83$, $SE = .02$) performance was significantly greater than sub-second performance ($M = .65$, $SE = .02$) on the most difficult ratio ($t(125) = -6.597$, $p < .001$). However, the opposite was true of the easiest ratio (Sub-Second: $M = .94$, $SE = .01$; Supra-Second: $M = .89$, $SE = .01$, $t(125) = 2.917$, $p = .004$). Sub- and supra- second timing was comparable on the two intermediate ratios, *p's* > .1, See Figure 4.7.

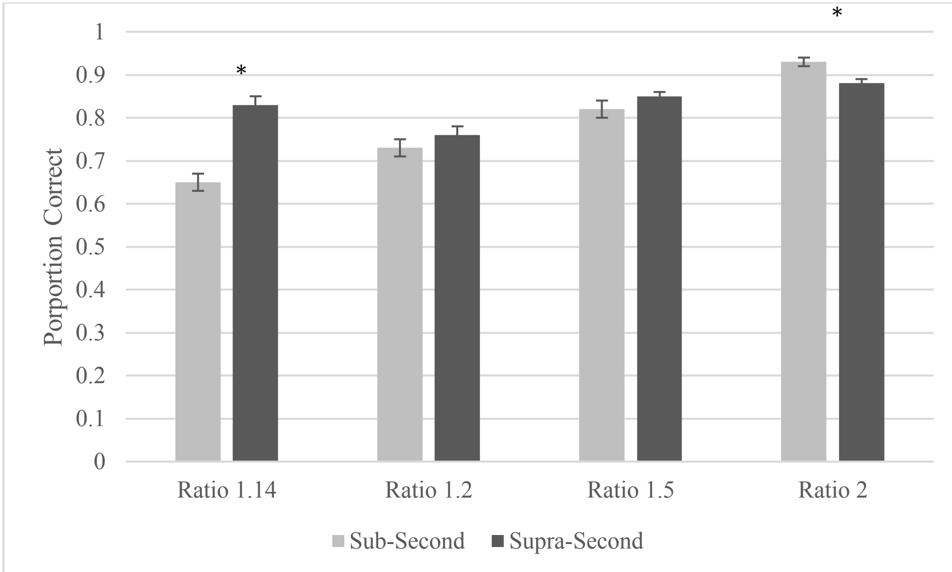

Figure 4.7. The effect of duration range appears to be driven by performance on the most difficult ratio comparison.

Next, we explored the effect of reproduction feedback on adults' temporal discrimination performance by conducting a 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANOVA on temporal discrimination performance. There was a main effect of Time, $F(1, 123) = 4405$, $p = .038$, $\eta_p^2 = .035$), indicating that performance was significantly better at post-test (M = .82, SE = .007) compared to pre-test (M = .81, SE = .007). Neither the main effect of Condition, $p = .918$, nor the Time x Condition interaction, $p = .166$, reached significance. See Figure 4.8[7].

[7] We again tested whether feedback impacted sub- and supra-second timing differently; however, neither ANOVA revealed a significant Time x Condition interaction, p's > .1.

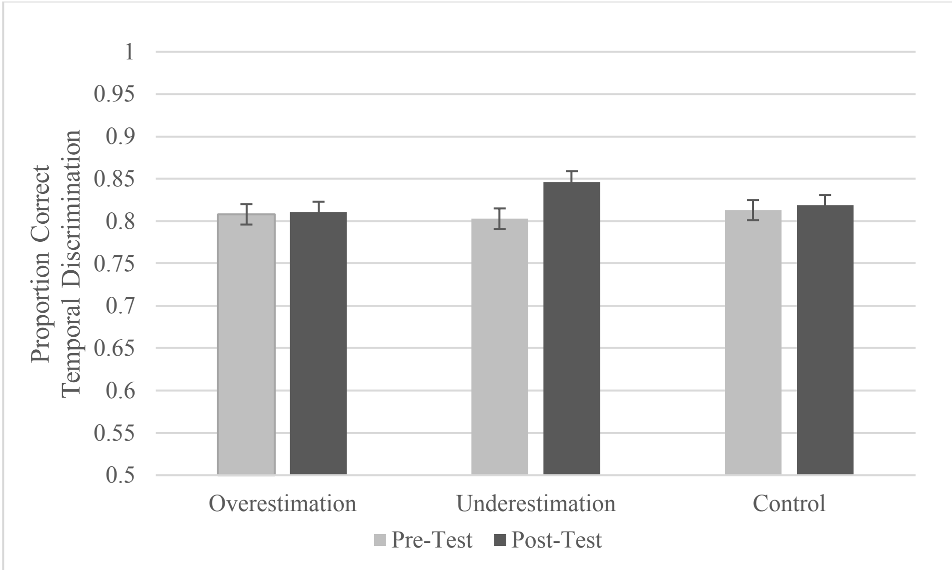

Figure 4.8. Performance improved after feedback in the underestimation condition, but not in the overestimation or control conditions.

Finally, we tested for correlations between changes in reproduction performance and temporal discrimination at post-test in the over- and underestimation conditions separately, to determine whether providing accurate feedback (i.e., underestimation feedback) resulted in better temporal discrimination performance. There was a significant correlation between change in average estimate and post-test temporal discrimination performance for those assigned to the overestimation condition only, $r = .475$, $p = .002$. That is, the degree to which participants in the overestimation condition adjusted their responses to be closer to the accurate duration was related to performance on the discrimination task at post-test. No other correlations reached significance, *p's* > .1.

Discussion

In Experiment 1, we found that shifting adults' symbolic mapping did not subsequently impact their performance on a nonsymbolic timing task. However, we wanted to ensure that this effect did not reflect difficulty of transferring the mapping from a symbolic task to a nonsymbolic one. Additionally, we wanted to determine whether individuals were able to calibrate their own internal clock according to the feedback provided. As in Experiment 1, adults attended to the feedback and modified their performance during the initial reproduction task: participants assigned to the overestimation condition again adjusted their slopes proportionally, whereas those in the underestimation and control conditions adjusted their estimates in a constant fashion. Although participants received feedback and thus adjusted their estimates on a nonsymbolic temporal reproduction task, performance on the subsequent temporal discrimination task was unaffected. While we found a positive correlation between the amount of change from pre- to post-test on the temporal reproduction task and post-test temporal discrimination performance, this appeared to be driven by those in the overestimation condition. Thus, counter to our predictions, making participants estimates more precise (as was true of the underestimation condition) did not lead to enhanced temporal discrimination performance.

General Discussion

We began this study with two aims: 1) to determine whether a correlation exists between symbolic and nonsymbolic timing in adults, and 2) to determine whether shifting adults' symbolic mapping of time subsequently impacted their nonsymbolic temporal discrimination performance. In addition, we tested whether using feedback to calibrate one's internal clock subsequently impacted nonsymbolic temporal discrimination performance. Across both experiments, adults attended to the feedback provided and adjusted their estimates accordingly. Counter to our predictions, however, we found no correlation between adults symbolic and nonsymbolic timing abilities, nor did adjusting adults' symbolic or nonsymbolic mapping impact performance on a subsequent nonsymbolic task.

Relation between Symbolic and Nonsymbolic Timing

Despite recent working showing a correlation between children's symbolic and nonsymbolic timing (Hamamouche & Cordes, accepted), this relation does not hold in adulthood. Given comparable findings in the numerical cognition literature, our study joins others questioning the strength of the relation between nonsymbolic and symbolic quantity processing beyond childhood (Inglis et al., 2011; Price, Palmer, Battista, & Ansari, 2012). While there was no correlation between symbolic and nonsymbolic timing, Experiment 2 replicated prior work showing a strong correlation between temporal reproduction and temporal discrimination in adults (see Cicchini et al., 2012).

Effects of Adjusting Mapping Using Feedback

In the present study, we used feedback in order to modify participants' symbolic and nonsymbolic mapping of time. Importantly, participants attended to feedback and adjusted their estimates accordingly: Individuals who received overestimation feedback lowered their estimates, whereas those in the underestimation condition raised their estimates. Given that we found this pattern to hold across both experiments, it seems likely that our feedback manipulation was convincing. However, we found that the way in which participants adjusted their feedback depended upon the feedback received. Across both experiments, receiving underestimation feedback led to a constant shift in temporal estimates, whereas receiving overestimation feedback led to a proportional shift in temporal estimates. Given that those in the overestimation condition only had so much room to reduce their estimates, it is likely that they only adjusted their estimates of longer durations that had more room for adjustment. Those in the underestimation condition however, did not have this issue as they could increase their estimates regardless of the duration, resulting in a constant shift.

Our main question was determining whether shifting one's symbolic or nonsymbolic mapping of time subsequently impacted performance on a nonsymbolic task. We predicted that adjusting the mapping would impact temporal discrimination in one of two possible ways: 1) adjusting estimates constantly may impact subsequent temporal discrimination performance, or 2) accurately adjusting estimates through underestimation feedback may improve temporal discrimination performance. Despite these predictions, neither experiment revealed a significant Time x Condition interaction, nor was there a unique relation between change in temporal discrimination performance and change in temporal estimation in the underestimation condition. Taken together with the lack of correlation between nonsymbolic and symbolic timing, it seems like these two abilities may be less related in adulthood. Moreover, these data appear to undercut the refinement hypothesis, suggesting that one's symbolic mapping does not impact one's nonsymbolic abilities. However, it remains possible that children who are still in the process of learning temporal symbols may show evidence of the refinement hypothesis, such that learning symbols shapes nonsymbolic abilities earlier in development. Future work with children will be critical for testing this possibility.

Conclusions

Together, these data suggest that while temporal estimates are malleable given feedback, temporal discrimination performance remained unaffected. This indicates that symbolic and nonsymbolic timing may be less related in adulthood. Given that these abilities seem more closely linked in childhood, future research will be important for 1) assessing the strength of this relation in childhood and 2) determining whether shifting children's symbolic mapping subsequently impact nonsymbolic temporal abilities.

Appendix. Covariate Analyses

We also conducted the mixed measures ANOVAs including pre-test performance as a covariate. By doing so, we were able to test the effect of pre-test performance on Time and Condition separately. However, because entering the covariate adjusts pre-test performance in all conditions to be equivalent, we must cautiously interpret any Time x Condition interactions found as they do not necessarily reflect accurate changes in performance.

Experiment 1

When including pre-test average estimate as a covariate, the 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANCOVA on average estimate revealed a main effect of Time, $F(1,126) = 4.975$, $p = .027$, $\eta_p^2 = .038$, such that estimates were longer at post-test (M = 838.316, SE = 22.064) compared to pre-test (M = 829.425, SE = .000). There was also a main effect of Condition, $F(2, 126) = 49.151$, $p < .001$, $\eta_p^2 = .438$, that was qualified by a significant Time x Condition interaction, $F(2, 126) = 49.151$, $p < .001$, , $\eta_p^2 = .438$. Follow-up analyses revealed that those in the overestimation condition significantly reduced their estimates (Pre-Test: M = 829.425, SE = .000; Post-Test: M = 609.333, SE = 39.498, $p < .001$) and those in the underestimation condition significantly increased their estimates (Pre-Test: M = 829.425, SE = .000; Post-Test: M = 1135.46, SE = 37.94, $p < .001$). Participants assigned to the control condition however, showed no changes in their estimates (Pre-Test: M = 829.425, SE = .000; Post-Test: M = 770.16, SE = 38.355; $p = .125$), See Figure A4.1. Finally, there was a significant Time x Covariate interaction, $F(1, 126) = 8.137$, $p = .005$, $\eta_p^2 = .061$. A negative correlation between pre-test average estimate and change in average estimate ($r = -.254$, $p = .004$) suggested that participants with greater estimates at pre-test reduced their estimates more than those who started out with lower estimates, See Figure A4.2.

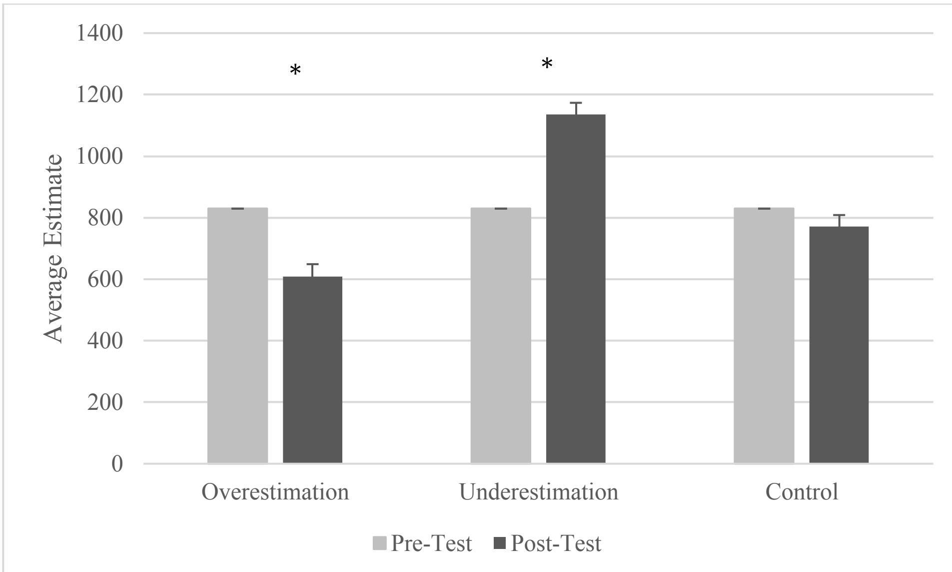

Figure A4.1. Participants adjusted their estimates according to the feedback provided.

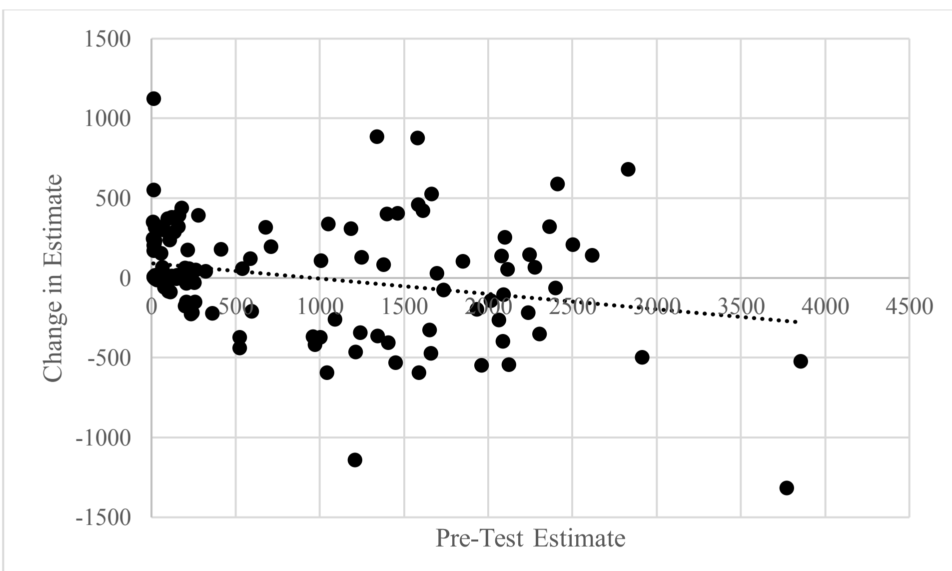

Figure A4.2. Participants who provided a larger estimate at pre-test showed a greater reduction in their estimates over time.

When including pre-test slope as a covariate, the 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANCOVA on slope showed a main effect of Condition, $F(2, 125) = 3.231$, $p = .043$, $\eta_p^2 = .049$, that was qualified by Time x Condition interaction, $F(2, 125) = 3.231$, $p = .043$, $\eta_p^2 = .049$. Follow-up analyses revealed that those in the overestimation and control conditions had significantly shallower slopes at post-test (Overestimation: Pre-Test: $M = .485$, $SE = .000$; Post-Test: $M = .417$, $SE = .034$, $p = .044$; Control: Pre-Test: $M = .485$, $SE = .000$; Post-Test: $M = .392$, $SE = .032$, $p = .005$), whereas those in the underestimation condition had comparable slopes at both time points (Pre-Test: $M = .485$, $SE = .000$; Post-Test: $M = .503$, $SE = .033$, $p = .593$), See Figure A4.3. There was also a significant Time x Covariate interaction, $F(1, 125) = 27.154$, $p < .001$, $\eta_p^2 = .178$. A negative correlation between pre-test average slope and the amount of change in slope over time ($r = -.428$, $p < .001$), suggested that participants with steeper slope at pre-test showed a greater reduction in slope compared to participants that started with shallower slopes, See Figure A4.4.

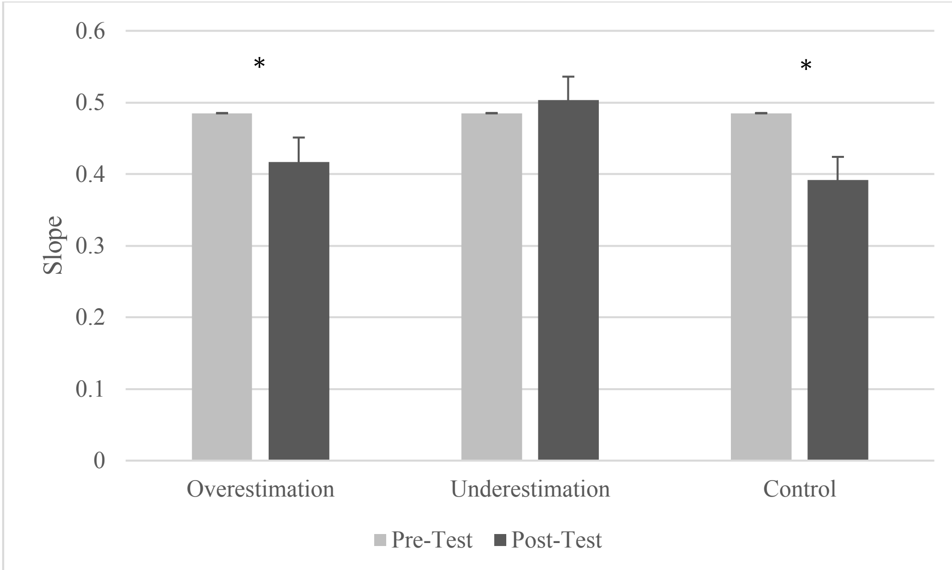

Figure A4.3. Slopes were shallower at post-test for those in the overestimation and control conditions.

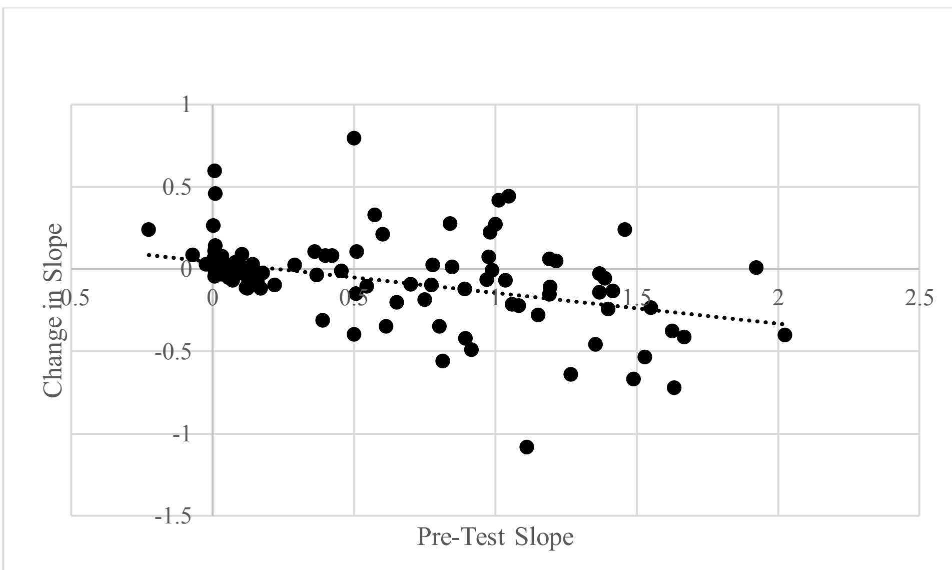

Figure A4.4. Participants with a steeper slope at pre-test showed a greater reduction in slope over time.

When including pre-test temporal discrimination as a covariate, the 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANCOVA on temporal discrimination performance revealed a main effect of Time, $F(1, 127) = 80.081$, $p < .001$, $\eta_p^2 = .387$, such that performance was greater at post-test (M = .819, SE = .006) compared to pre-test (M = .797, SE = .000). There was a marginal effect of Condition, $F(2, 127) = 2.980$, $p = .054$, $\eta_p^2 = .045$, that was qualified by a marginal Time x Condition interaction, $F(2, 127) = 2.980$, $p = .054$, $\eta_p^2 = .045$. Performance was greater at post-test for those in the underestimation condition (Pre-Test: M = .797, SE =.000; Post-Test: M = .840, SE =.011, $p < .001$). However, those in the overestimation (Pre-Test: M = .797, SE = .000; Post-Test: M = .815, SE = .011, $p = .116$) and control conditions (Pre-Test: M = .797, SE =.000; Post-Test: M = .802, SE =.011, $p = .672$) had comparable performance at pre- and post-test, See Figure A4.5. While this shows our predicted pattern of results, we must be careful when interpreting this interaction as the model adjusts pre-test performance to be identical across conditions. There was also a significant Time x Covariate interaction, $F(1, 127) = 75.518$, $p < .001$, $\eta_p^2 = .373$. We again followed up on the Time x Covariate interaction by conducting a correlational between pre-test temporal discrimination and the amount of change on the temporal discrimination task (post-test – pre-test). A strong negative correlation ($r = -.592$, $p < .001$) indicated that participants who had worse performance at pre-test showed a larger amount of change on the temporal discrimination than participants who had better pre-test performance, See Figure A4.6.

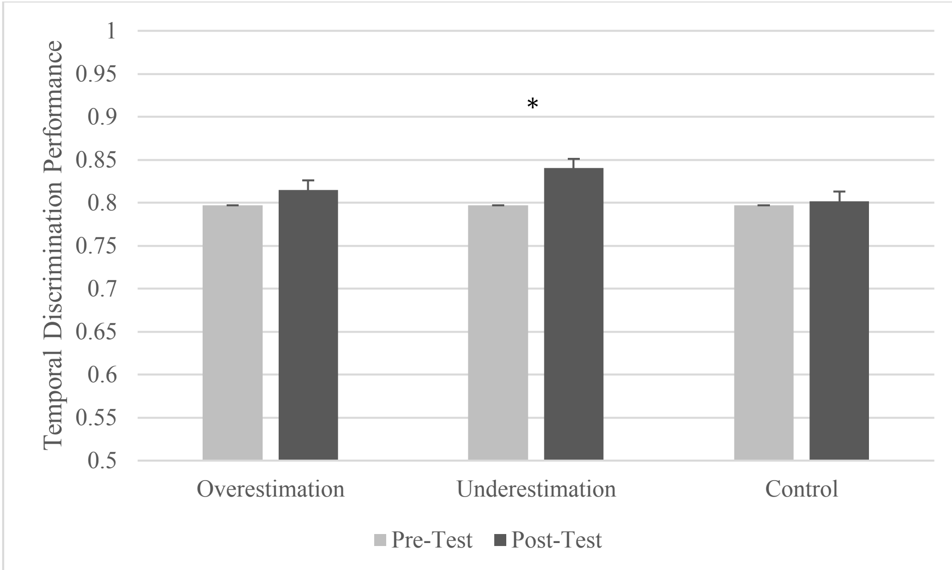

Figure A4.5. Participants in the underestimation condition showed improved temporal discrimination performance at post-test.

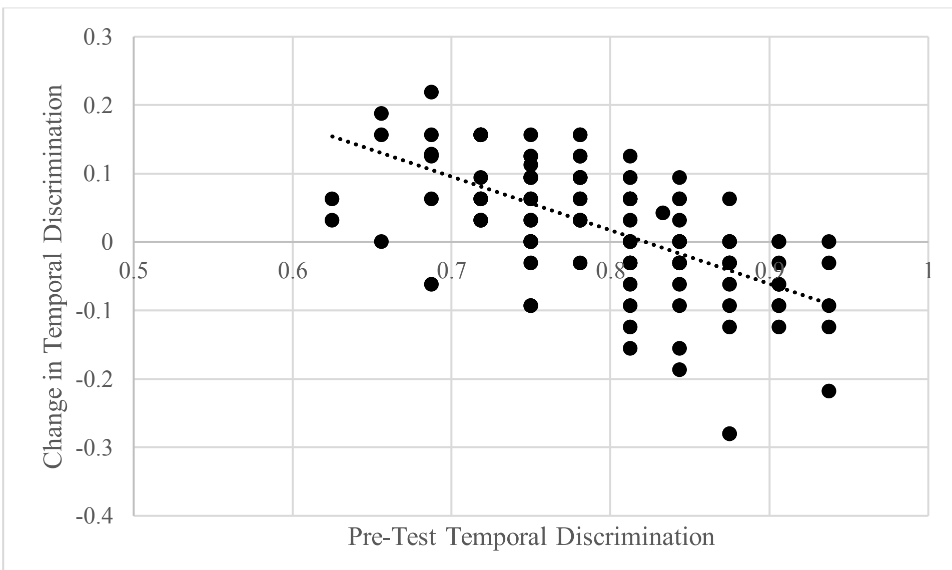

Figure A4.6. Participants with worse pre-test performance showed the greatest change in their performance over time.

Experiment 2

When including pre-test estimate as a covariate, the 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANCOVA on average estimate showed a main effect of Time, $F(1, 120) = 10.254$, $p = .002$, $\eta_p^2 = .067$, such that estimates were on average longer at post-test (M = 1366.95, SE = 15.938) compared to pre-test (M = 1353.402, SE = .000). There was also a main effect of Condition, $F(2, 120) = 38.699$, $p < .001$, $\eta_p^2 = .392$, that was qualified by a significant Time x Condition interaction, $F(2, 120) = 38.699$, $p < .001$, $\eta_p^2 = .392$. Follow-up analyses revealed that those in the overestimation condition significantly decreased their estimates at post-test (Pre-Test: M = 1353.402, SE = .000; Post-Test: M = 1192.343, SE = 27.783; $p < .001$), whereas those in the underestimation significantly increased their estimates at post-test (Pre-Test: M = 1353.402, SE = .000; Post-Test: M = 1537.098, SE = 27.717; $p < .001$). Those in the control condition had comparable performance at pre- and post-test (Pre-Test: M = 1353.402, SE = .000; Post-Test: M = 1371.516, SE = 27.412; $p = .510$), See Figure A4.7. There was also a significant Time x Covariate interaction, $F(1, 120) = 7.894$, $p = .006$, $\eta_p^2 = .067$. A negative correlation between pre-test average estimate and change in average estimate ($r = -.181$, $p = .044$) revealed that participants with shorter estimates at pre-test were more likely to increase their estimates over time, whereas those with longer estimates at pre-test were likely to decrease their estimates over time, See Figure A4.8.

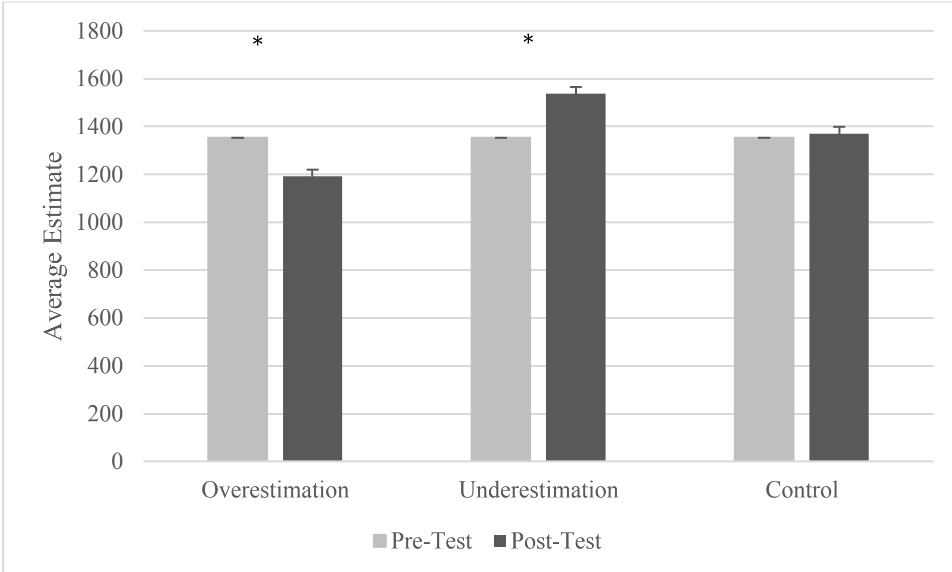

Figure A4.7. Participants attended to feedback and adjusted their estimates accordingly.

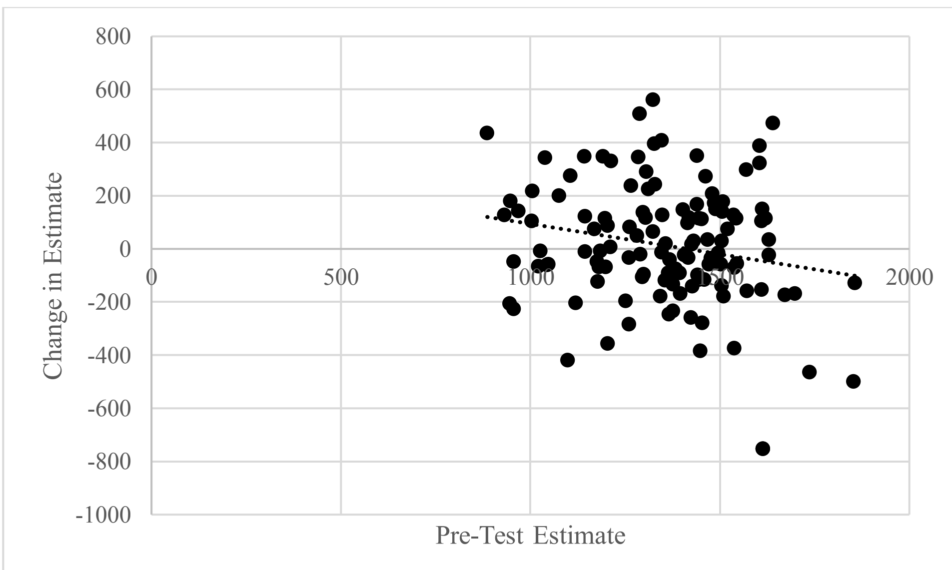

Figure A4.8. Participants with the shortest estimates at pre-test were more likely to increase their estimates, whereas participants with larger estimates at pre-test were more likely to decrease their estimates.

When including pre-test slope as a covariate, the 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANCOVA on slope showed a main effect of Time, $F(1, 118) = 17.495$, $p < .001$, $\eta_p^2 = .129$, such that slope was shallower at post-test ($M = .708$, $SE = .020$) compared to pre-test ($M = .789$; SE

= .000). There was also a main effect of Condition $F(2, 118) = 3.224$, $p = .043$, $\eta_p^2 = .052$ that was qualified by a significant Time x Condition interaction, $F(2, 118) = 3.224$, $p = .043$, $\eta_p^2 = .052$. Those in the overestimation condition had greater slopes at pre-test ($M = .789$, $SE = .000$) compared to post-test ($M = .637$, $SE = .035$, $p < .001$), whereas those in the underestimation (Pre-Test: $M = .789$, $SE = .000$; Post-Test: $M = .757$, $SE = .034$, $p = .357$) and control (Pre-Test: $M = .789$, $SE = .000$; Post-Test: $M = .730$, $SE = .035$, $p = .095$) conditions had comparable slopes across time, See Figure A4.9. Finally, there was a significant Time x Covariate interaction, $F(2, 119) = 22.416$, $p < .001$, $\eta_p^2 = .159$. A negative correlation between pre-test slope and change in slope ($r = -.378$, $p < .001$) suggested that participants with a steeper slope at pre-test tended to show a reduction in slope over time, See Figure A4.10.

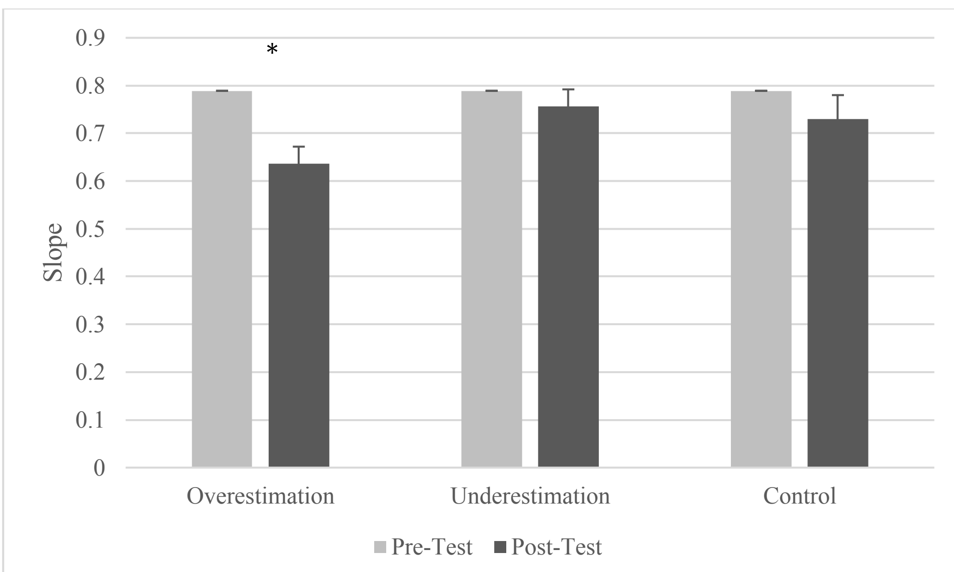

Figure A4.9. Participants in the overestimation had shallower slopes at post-test.

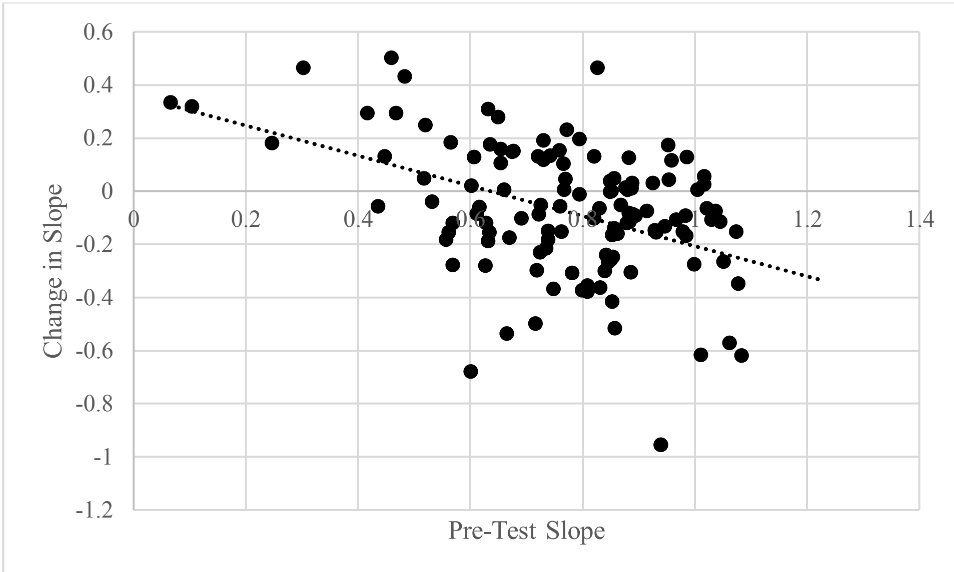

Figure A4.10. Participants with shallower slopes at pre-test tended to have a positive change in slope, whereas those who began with a steeper slope tended to have a negative change.

When including pre-test temporal discrimination performance as a covariate, the 2 (Time: Pre- versus Post-Test) x 3 (Condition: Overestimation, Underestimation, Control) mixed measures ANCOVA revealed a main effect of Time, $F(1, 122) = 39.071$, $p < .001$, $\eta_p^2 = .243$, such that performance was better at post-test ($M = .821$, $SE = .007$) compared to pre-test ($M = .806$, $SE = .000$). There was also a significant Time x Covariate interaction, $F(1, 122) = 36.071$, $p < .001$, $\eta_p^2 = .243$. A negative correlation between pre-test temporal discrimination and change in temporal discrimination ($r = -.499$, $p < .001$) suggested that individuals who had better performance at pre-test did not show as much growth as those who started with lower performance, See Figure A4.11. No other main effects or interactions reached significance, *p's* > .3.

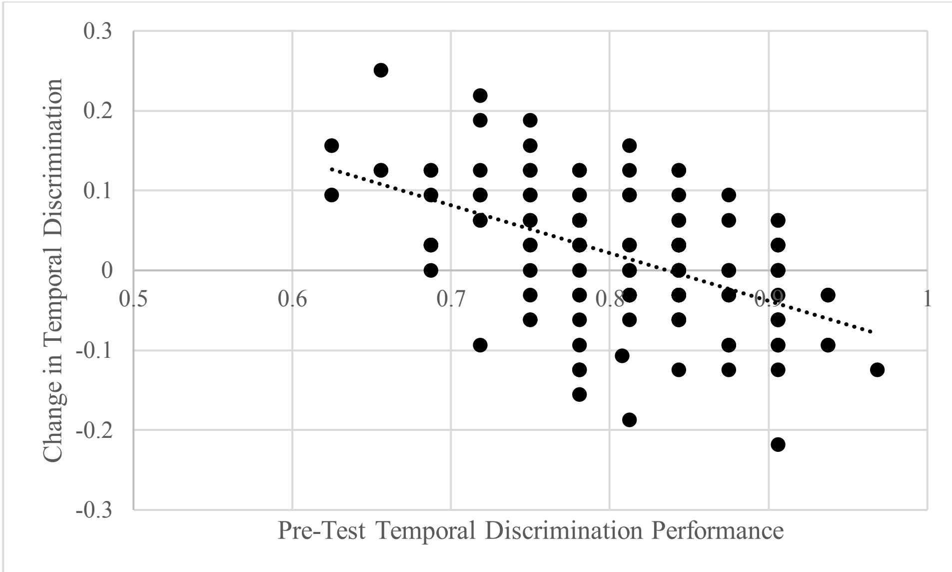

Figure A4.11. Participants with worse performance at pre-test showed greater improvements over time.

CHAPTER 5: MECHANISMS UNDERLING THE REFINEMENT HYPOTHESIS: THE CASE OF MEMORY AND ATTENTION

Abstract

Studies suggest that learning symbols for quantity is related to our ability to track that quantity (e.g., Mussolin, Nys, Content, & Leybaert, 2014; Shusterman et al., 2016). These investigations however, have been correlational in nature, leaving it unknown whether experimental manipulations may reveal a causal relation between acquiring formal symbols and perceiving quantity. In the present study, we investigate whether learning symbols to represent different surface area sizes impacts adults' ability to discriminate surface area (in a nonsymbolic manner). We also tested potential mechanisms underlying these changes: that is, whether heightened attention and/or led to enhanced memory encoding of spatial dimensions may underlie changes in spatial acuity. Adult participants (N = 126) completed either a training in which they learned symbols to represent surface area (symbol condition), color (control condition), or attended to surface area without learning labels (control condition). Before and after training, participants also completed area discrimination, area salience, and memory encoding tasks. Performance on the symbolic training task was correlated with changes in area discrimination performance and memory encoding, but only for those who learned symbols to represent surface area. This suggests that learning symbols to represent surface areas may be related to nonsymbolic area discrimination performance, and memory encoding may play a role in these changes. These findings, however, do not provide support for a causal relation. We end by identifying areas of future research that will be fruitful for further understanding the relation between symbol acquisition and basic quantity processing.

Learning (verbal or written) symbols to represent quantities is a critical component of early education. For example, preschool classrooms focus on counting and early math concepts, whereas elementary age students are taught temporal units of measurement (e.g., a minute is longer than a second; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). While numerous studies have revealed that one's understanding of symbols is related to their ability to track quantity (e.g., Bonny & Lourenco, 2014; Halberda, Mazzocco, & Feigenson, 2008; Hamamouche & Cordes, accepted; Inglis, Attridge, Batchelor, & Gilmore, 2011; Libertus, Feigenson, & Halberda, 2011; Libertus, Odic, & Halberda, 2012; Mazzocco, Feigenson, & Halberda, 2008; Odic, Lisboa, Eisinger, Olivera, Maiche, & Halberda, 2016; Starr, Libertus, & Brannon, 2013; for meta-analysis see: Schneider, Beeres, Coban, Merz, Schmidt, Stricker, de Smedt, 2017), the exact effect of learning symbols on one's ability to track quantities is unknown. That is, very few studies have used experimental manipulations to test whether learning quantity symbols directly refines one's ability to perceive quantities (although see Lindskog, Winman, & Poom, 2016). In the current study, we investigate whether learning symbols to represent surface area affects adults' spatial acuity.

Relations between Quantity Representations: The Case of Number and Time

Substantial evidence suggests that we have two distinct systems for representing quantity, symbolic and nonsymbolic systems (see Feigenson, Dehaene, & Spelke, 2004; Halberda, Feigenson, & Mazzocco, 2008). Our nonsymbolic representations of quantities are approximate and independent of language. For example, our ability to estimate which of two plates has a greater number of cookies or which of two cookies is bigger relies upon nonsymbolic quantitative abilities. Symbolic representations, on the other hand, are dependent upon language and include verbal labels or signs, discrete gestures, or written symbols which allow for an exact system of measurement (e.g., Arabic numerals for number, units of measurement for time, etc.). Several studies have found a relation between nonsymbolic and symbolic quantitative skills (see Schneider et al., 2017). For example, fourteen-year olds' performance on a nonsymbolic numerical discrimination task predicts their math SAT performance (Halberda et al., 2008) and understanding of temporal units of measurement is related to nonsymbolic timing abilities in elementary-aged children (Hamamouche & Cordes, accepted).

Not only do correlations exist between nonsymbolic and symbolic quantitative abilities, newer work suggests that this relation may be causal. Hyde, Khanum, & Spelke (2014) showed that training on a nonsymbolic numerical task led to enhanced math abilities in children, suggesting that nonsymbolic abilities may form the foundation for learning about symbolic representations. While this points to a causal relation between nonsymbolic and symbolic abilities, it remains unknown whether this relation is bidirectional, such that learning symbols might also refine nonsymoblic acuity. However, a few studies hint at this possibility. Cross-lagged longitudinal studies show that children's symbolic understanding of number predicts their later nonsymbolic numerical acuity (Lyons, Bugden, Zheng, De Jesus, & Ansari, 2018; Matejko & Ansari, 2014; Mussolin, Nys, Content, & Leybaert, 2014). However, none of these studies involved experimental manipulations. In a similar investigation, young children's nonsymbolic representations of number – as assessed by a numerical discrimination task – were tracked over a period of six months as they became proficient counters. As soon as children acquired the cardinal principle – indicative of mastery of verbal counting, a symbolic representation of number - their nonsymbolic numerical acuity drastically improved (Shusterman, Slusser, Halberda, & Odic, 2016). The authors concluded that learning symbols for quantities refines one's nonsymbolic representations of that quantity; however, it is also possible that a drastic shift in their numerical acuity led to learning the counting principle. While this finding is suggestive of a causal relation, only one study has intentionally manipulated symbol learning to explore this relation (Lindskog et al., 2016). Thus, the current study serves as one of the first investigations of the effect of symbol acquisition on nonsymbolic quantity processing.

Why Symbols Shape our Perception: Evidence from Language

In addition to testing whether learning symbols shapes nonsymbolic representations of time, it is also important to determine the mechanism(s) underlying this change. As such, the second aim of this study was to assess the mechanism(s) underlying improved nonsymbolic acuity following symbol acquisition. *Why* might the acquisition of quantity symbols impact nonsymbolic representations of quantity? Lera Boroditsky – a cognitive scientist who explores how language shapes cognition – argues that, "Languages force us to attend to certain aspects of our experience by making them grammatically obligatory. Therefore, speakers of different languages may be *biased to attend* to and *encode* different aspects of their experience while speaking." (Boroditsky, 2001, p. 2; stress added). That is, it is possible that acquiring symbolic knowledge may (1) guide attention towards relevant cues and/or (2) facilitate memory encoding of relevant information. The present study attempts to address each of these hypotheses by assessing whether acquisition of spatial symbols heightens attention to and/or enhances encoding of spatial extent.

Heightened Attention. A substantial literature suggests that the language we learn may guide our attention to different aspects of the world. For instance, hearing a verbal label has been shown to direct individuals' attention towards previously unseen objects at the perceptual level (Lupyan & Ward, 2013). Importantly, this effect is unique to hearing a verbal label (e.g. "frog"), but not a related, nonverbal sound (e.g., "ribbet"; Boutonnet & Lupyan, 2015). Illustrating this effect, Ostarek and Huettig (2017) showed participants a binocular fusion, such that each eye saw a different image - one eye saw a visual mask and the other eye saw an image (e.g., a bottle, banana, palm tree, etc.), which when looking at simultaneously resulted in the images appearing very abstract. Prior to the onset of the images, participants heard either a congruent word (e.g. "bottle" if one eye would subsequently see a bottle in one eye) or an incongruent word (e.g., "banana" if one eye would then see a bottle in one eye). After seeing the images, participants were asked whether or not they saw that item. Performance was enhanced on congruent trials, compared to incongruent ones (Ostarek & Huettig, 2017), indicating that hearing labels guided participant's attention towards important aspects in their environment.

If learning area symbols guides attention towards the relevant features, then in our study, participants' attention should be guided towards spatial information and they should be better able to discriminate subtle differences in surface area. In order to assess this, participants completed an irrelevant task, tapping their numerical discrimination abilities (subsequently referred to as the area salience task), in which they were asked to compare the relative numerosity of two arrays of dots. Importantly, the surface area of the enumerated dots in the task were manipulated such that half of the trials were congruent in surface area (the average size of the dots was the same in both arrays, such that the cumulative area of the more numerous array was also be greater) and half of the trials were incongruent (cumulative surface area remained the same regardless of the numerosity).

Prior research reveals adults typically perform better on congruent compared to incongruent trials, revealing a slight reliance on cumulative area during the numerical task (Clayton & Gilmore, 2015; DeWind & Brannon, 2012; Gilmore et al., 2013; Hurewitz et al., 2006; Nys & Content, 2012). If learning area symbols heightens one's attention to surface area, then it is predicted that participants will show a greater attention to surface area of the dot arrays during the incongruent trials, resulting in *worse* performance on incongruent trials (and potentially better performance on congruent trials) at post-test.

Memory Encoding. Memory encoding for surface area may also be enhanced after acquiring symbols. In line with this hypothesis is work showing that one's language often determines how information is encoded (Fausey & Boroditsky, 2011). In one study, English and Spanish speakers viewed intentional or accidental events. Importantly, while the English language uses agentive expressions regardless of whether an event was intentional or accidental, Spanish use non-agentive expressions when describing accidental expressions. While speakers of both languages remembered the agents involved in intentional events equally well, English speakers recalled the agents involved in accidental events more accurately than Spanish speakers (Fausey & Boroditsky, 2011). Researchers concluded that the English speakers encoded information about the agents in the accidental events, suggesting that their language impacted encoding and memory for events. Relatedly, other work shows that bilingual speakers remember information differently depending on the language in which the information was encoded (Schroeder & Marian, 2014). If it is the case that language may impact memory encoding, then we would expect participants who learn labels for distinct surface areas may be at an advantage for encoding surface area information.

To test whether learning quantity symbols leads to enhanced memory encoding, participants also completed a memory encoding task before and after learning symbols to represent surface area. If learning symbols leads to better memory encoding of area, one would expect enhanced performance on the memory encoding task at post-test, only for those participants who learned symbols to represent surface area.

The Current Study

In the current study, participants were randomly assigned to one of three different training conditions: (1) learning novel labels to represent distinct measurements of surface area, (2) learning labels to represent color (control condition), or (3) exposure to surface area without learning labels (control condition). Participants' nonsymbolic spatial acuity was assessed both before and after symbol acquisition. Participants also completed an area salience task and memory task allowing us to explore how acquisition of quantitative symbols shapes nonsymbolic representations of quantity through either heightened attention to surface area and/or enhanced memory encoding of surface area information.

Methods

Participants

Undergraduates (N = 126, 102 females, M_{age} = 18.77 years) participated for course credit. Required sample size was determined a priori using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009). An additional 8 participants completed the study, but were excluded for performing below chance on three or more tasks (n = 2), or due to computer error (n = 6).

Stimuli

The same set of 42 irregular shape stimuli was used for both the memory and training tasks. There were 16 versions of each shape that varied by four different sizes (500 mm^2, 750 mm^2, 1120 mm^2, 1680 mm^2) and four different shades of blue (saturation: 53%, 61%, 70%, and 80%). That is, for each size (e.g., 500 mm^2), we had four individual shapes in each of the four colors. We did this so that regardless of training condition, all participants would see the same shapes. Pilot testing revealed that color discrimination precision was higher than that of area discrimination precision, thus the hues varied by a more difficult 1.15 ratio compared to the 1.5 ratio change in surface areas. The ratio between each color/item size was consistent across the four dimensions. In other words, there was a 1.15 ratio difference between the lightest shade and the next lightest shade, the second lightest shade and the third shade, etc.

Procedure

After consent was obtained, the experimenter took the participant into a quiet room. Participants were randomly assigned to one of three training conditions (42/condition). Participants completed the area discrimination, area salience, and memory task both before (pre-test) and after (post-test) participating in training:

(1) Area discrimination (adapted from Odic, et al., 2013; Pre-test and Post-test)

Participants viewed two irregular shapes (one blue and the other yellow) presented on a light gray background (See Figure 5.1). The shapes were displayed briefly (1500 ms, as per Odic et al., 2013) and the participants were asked to select which shape was larger in surface area. Participants made their response by pressing the left arrow key if the yellow shape was larger and the right arrow key if the blue shape was larger. Identical to Odic et al. (2013), the ratios between each shape differed by 1.14, 1.2, 1.5, 2.0, and 3.0 and each ratio was displayed sixteen times for a total of 80 trials. On each trial, one of the irregular shapes contained a surface area (500 mm^2, 750 mm^2, 1120 mm^2, 1680 mm^2) on which participants would later be trained.

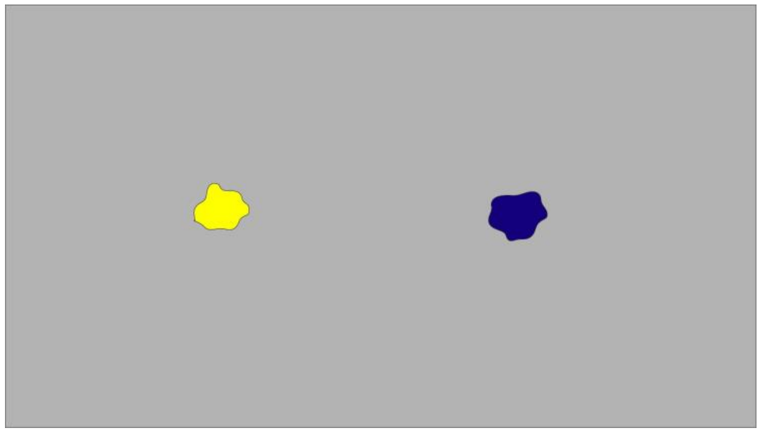

Figure 5.1. Example of area discrimination trial.

(2) Area salience task (Panamath, Halberda et al., 2008; Pre-test and Post-test)

Participants then completed an Area Salience task using the Panamath Program (Halberda et al., 2008). The experimenter told participants they would see two dot arrays and they were to decide which array was more numerous. Because Panamath changes the difficulty of trials based on age, to maintain consistency across all subjects, the age was set to 18 and the duration was set to 5 minutes. Given these settings, participants completed 80 trials. Importantly, half of the trials were congruent, such that cumulative surface area and numerosity were positively correlated. In the other half of the trials, cumulative surface area remained consistent across both arrays, such that the more numerous array contained smaller individual dots (i.e., incongruent trials).

(3) Memory task (Pre-test and Post-test)

The memory task was a delayed match to sample task used to measure memory encoding. During the memory task, participants saw a target shape displayed for a brief period of time (1000 ms). Then, a visual mask appeared on the screen for 2000 ms. After the visual mask disappeared, two potential matches were visible on the screen and the participant was told to indicate which of the two stimuli matched the target stimulus in terms of surface area (See Figure 5.2 for a progression of the task). The three shapes were always the same color, but each shape was different, and the incorrect match only differed in surface area. The ratio differences between the two choices varied between 1.5, 2.24, and 3.36). Participants completed 50 trials. Although the memory task was very similar to the training, it differed in two important ways. First, during the memory task, the sample image disappeared, whereas the sample remained on the screen during training. Secondly, the delay between the match appearing and the choices appearing was longer during the memory task compared to training.

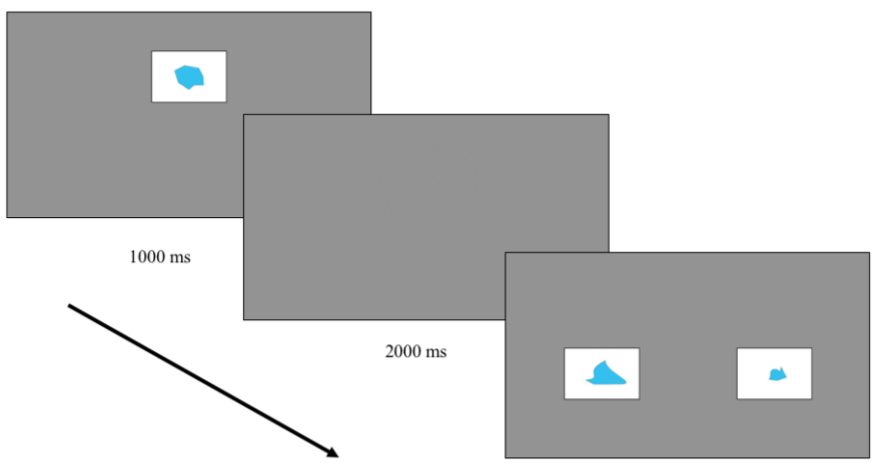

Figure 5.2. Progression of the memory task.

(4) Training

Participants were randomly assigned to one of three training conditions: Area Label, Nonsymbolic Area (control condition), or Color Label (control condition). In the Area Label condition, participants learned novel labels to represent distinct surface area measurements. In the Nonsymbolic Area condition, participants did not learn labels, but were exposed to the same exact stimuli as the two other conditions and were instructed to attend to the surface area of the items. In the Color Label Condition, participants learned novel labels for distinct shades of the color blue. We used novel labels rather than exact measurements to ensure that participants' previous knowledge did not influence their performance.

Area Label Training: At the beginning of the Area Label Training, the experimenter told the participant they were going to learn words to describe the surface areas of shapes. Then participants saw an image containing four different irregular shapes. Each shape was one of the training surface areas (500 mm^2, 750 mm^2, 1120 mm^2, or 1680 mm^2) and was in order from smallest (on the left) to the largest (on the right). The experimenter circled and labeled each shape as one dax, two daxes, three daxes, and four daxes, respectively. After the introduction to the shapes, the participants completed a match to sample task to learn the label. During the match to sample task, participants saw a sample shape at the top of the screen. Above the sample shape was the following statement: "The surface area of this shape is e.g., two daxes. Can you find the other shape with the surface area of two daxes?" After 1000 ms, two potential matches appeared at the bottom of the screen, and the participant had an unlimited amount of time to decide which of the two images matched the first image in terms of surface area. All three images (one sample, and two possible matches) remained on the screen until the participant made a response. Although the participant could take as long as they needed to decide, they were told to complete the task as quickly and accurately as possible. Participants had equal exposure to each of the four surface areas (e.g., 24 trials where the sample shape was 500 mm^2, 24 trials where the sample shape was 750 mm^2, …etc.), and the distractor shape was one of the other three sizes an equal number of times (e.g., if the sample shape was 500 mm^2, the distractor was 750 mm^2, 1120 mm^2, 1680 mm^2 the same number of times). Within each trial, the three shapes (sample and two potential matches) were the same shade of blue. However, across trials, participants received equal exposure to each of the shades of blue. This was to ensure equal exposure of color and size across conditions.

Nonsymbolic Area Training: Participants assigned to the Nonsymbolic Area condition completed an identical task, however, the surface areas were never labeled. When the participants viewed the image of the four shapes increasing in surface area, the experimenter circled each shape and said "This is the surface area of this shape." Then, on each match to sample trial, the statement: "This is the surface area of this shape. Can you find the other shape with the same surface area?" was written above the sample shape. Again, within trials, the three shapes were the same shade of blue, but across trials the shades of blue varied.

Color Label Training: Participants assigned to the Color Label Condition completed an identical task to the Area Label Condition; however, they were told they were learning new words to describe the color of the shapes. Thus, participants learned "one dax" represented the lightest shade of blue, "two daxes" represented the second lightest shade, and so forth such that "four daxes" represented the darkest shade of blue. On each trial, a statement prompted participants to find the shape with the specified color. In the Color Label Condition, within each trial, the sample and potential matches were all the same size; however, across trials, participants received equal exposure to the four different surface areas. Thus, within trials, the three shapes were identical in surface area, but across trials the size of the shapes changed so that participants in this condition still received a comparable amount of exposure to different sized shapes.

Across all three training conditions, trial by trial feedback was provided (a green check mark appeared for 1000 ms following a correct response, whereas a red X appeared for 1000 ms following an incorrect response). Participants completed one training block consisting of 96 trials.

After training, participants again completed the area discrimination task, area salience task, and the memory task (in that order).

Data Coding and Analyses

Accuracy (proportion correct) was the dependent variable for all tasks. We first excluded performance that was below chance (50%) on all tasks ($N_{\text{Memory Pre-Test}} = 7$, $N_{\text{Memory Post-Test}} = 1$). Then, any data points that were three standard deviations above or below the mean were excluded ($N_{\text{Area Salience Post-Test}} = 1$). We did not exclude training performance three standard deviations above/below the mean. For the area saliency task, we calculated a difference score (congruent trials performance – incongruent trial performance), thus a positive score would be indicative of better performance on the congruent trials, whereas a negative score would indicate better performance on the incongruent trials. One participants' difference score at pre-test was three standard deviations below the mean and so it was removed from analyses. No difference scores were three standard deviations above/below the mean were replaced.

Results

See Table 5.1 for average performance on each task and Table 5.2 for correlations between pre- and post-test performance. Performance on all tasks was significantly above chance (Area discrimination: Pre-test: $t(125) = 84.416$, $p < .001$; Post-Test: $t(125) = 72.915$, $p < .001$; Area Salience: Pre-test: $t(125) = 84.235$, $p < .001$; Post-test: $t(125) = 80.64$, $p < .001$; Memory: Pre-test: $t(118) = 24.419$, $p < .001$; Post-test: $t(124) = 26.238$, $p < .001$). Pre-test performance on the area discrimination did not correlate with pre-test performance on the area salience task ($r = .099$, $p = .271$) or the memory task ($r = .108$, $p = .244$). Significant accuracy ratio effects were found for both the area discrimination (Pre-Test: $F(4, 500) = 186.53$, $p < .001$, $\eta_p^2 = .602$; Post-test: $F(4, 500) = 157.32$, $p < .001$), $\eta_p^2 = .564$) and area salience tasks (Accuracy: Pre-test: $F(3, 375) = 378.16$ $p < .001$, $\eta_p^2 = .759$; Post-test: $F(3, 375) = 447.61$, $p < .001$, $\eta_p^2 = .782$).

	Area Discrimination	**Area Salience**	**Area Salience Difference Score**	**Memory**
Pre-test	.885 (.005)	.854 (.004)	.900 (.53)	.733 (.01)
Post-test	.874 (.005)	.853 (.004)	2.26 (.578)	.726 (.009)

Table 5.1. Average performance (standard error) on each task.

	Area Discrimination	**Area Salience**	**Area Salience Difference Score**	**Memory**
Area Labels	.674**	.622**	.187	.475*
Nonsymbolic Area	.607**	.404*	$.297^{+}$	.551**
Color Labels	.611**	.385*	.438*	.380*

Table 5.2. Correlations between pre- and post-test performance.
$^{**}p < .001$, $^{*}p < .01$, $^{+}p < .1$

Training Performance

Although the three training conditions were designed to be comparable, a one way ANOVA revealed marginal differences in accuracy on the training task, $F(2, 123) = 3.014$, $p = .053$. Follow-up independent sample t-tests revealed that participants assigned to the

Nonsymbolic Area (M = .884, SE = .008) condition outperformed those assigned to the Area Label condition (M = .845, SE = .017; $t(56.984) = -2.15$, p = .035). Training performance for those assigned to the Color Label Condition (M = .882, SE = .01) was marginally better than the Area Label condition ($t(75.26) = -1.78$, p = .078), but not different from the Nonsymbolic Area Condition ($t(82) = .172$, p = .864), See Figure 5.3.

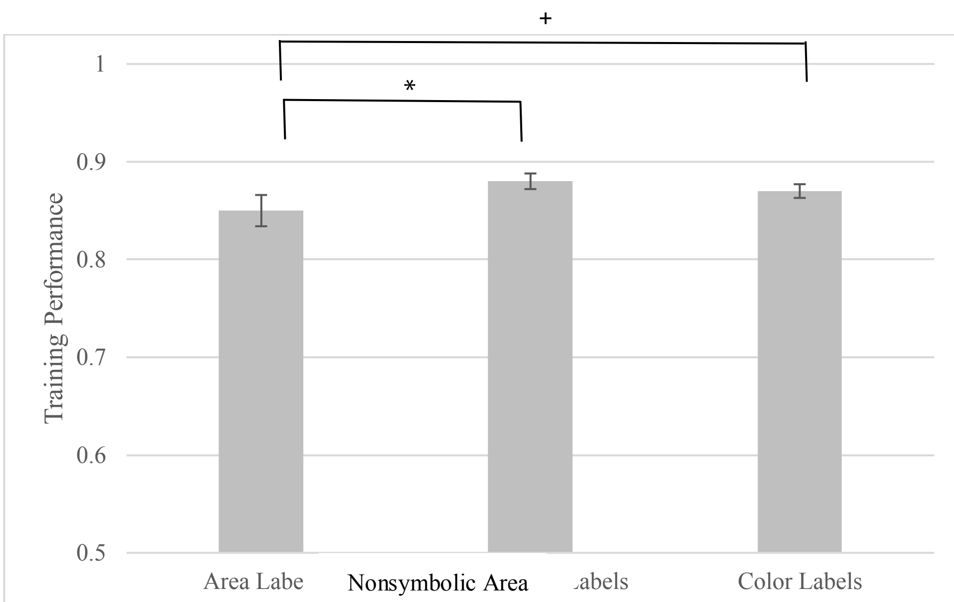

Figure 5.3. Performance on the training task.

Does learning symbols for surface area enhance nonsymbolic representations of area?

Area discrimination performance was comparable across the three conditions at pre-test, $F(2, 123) = 1.115$, p = .331. To test whether learning area symbols enhanced nonsymbolic representations of area, a 2 (Time: Pre versus Post Test) x 3 (Condition: Area Labels, Nonsymbolic Area, Color Labels) mixed measures ANOVA on area discrimination accuracy was conducted. There was a main effect of Time, $F(1, 123) = 6.852$, p = .010 , η_p^2= .053, such that area discrimination performance was worse at post-test (M = .874, SE

= .005) compared to pre-test (M = .885, SE = .005). No other main effects or interactions reached significance, p's > .1, See Figure 5.4.

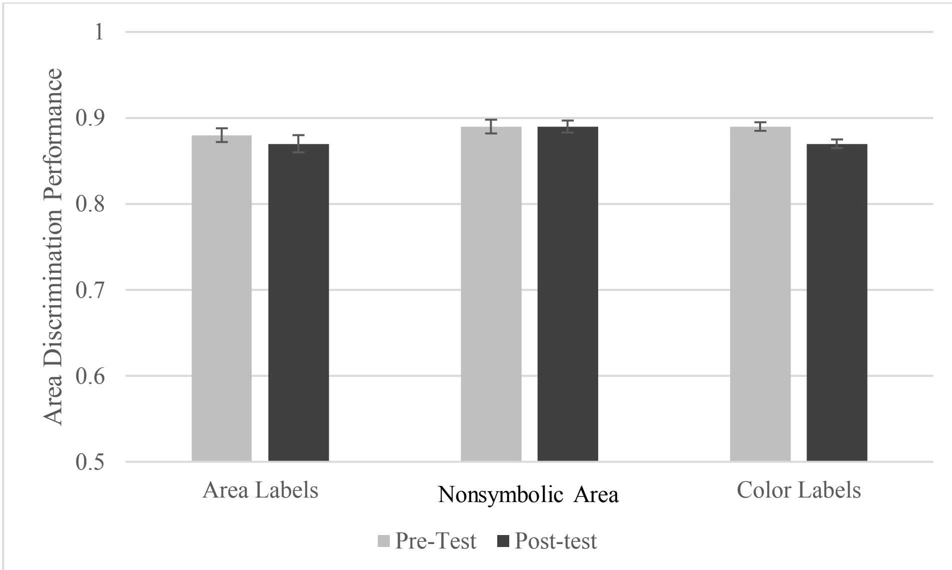

Figure 5.4. Learning area symbols did not impact nonsymbolic representations of space.

To further test the relation between our symbolic training and nonsymbolic discrimination task, we next conducted correlations between accuracy on the training task and improvement on the area discrimination task from pre- to post- (determined as the difference score between accuracy on the area discrimination task at post-test compared to pre-test in each condition separately). We controlled for pre-test performance to rule out individual differences in overall performance. A partial correlation (controlling for pre-test performance on the area discrimination task) revealed that accuracy on the Area Label training correlated with the change in area discrimination performance (r = .399, p = .010), suggesting that those participants who better learned area labels showed greater improvements in area discrimination performance. In contrast, there was no correlation between accuracy on the training and the area discrimination difference score (when controlling for pre-test performance, for the Nonsymbolic Area (r = .207, p =.195) or Color

Label (r = -.148, p = .356) conditions. While the correlation in the Area Labels was significantly stronger than that of the Color Labels (z = 2.52, p = .005), the difference between the Area Labels and Nonsymbolic Area conditions was not significant (z = 0.94, p = 0.17). Performance was marginally different between the Nonsymbolic Area and Color Labels conditions (z = 1.3, p = 0.055).

What mechanism(s) might lead to enhanced nonsymbolic representations of area?

Heightened Attention. One way ANOVAs revealed comparable performance among all three conditions at pre-test (Area salience: $F(2, 123) = 1.306$, p = .275; Area salience difference score: $F(2, 122) = .382$, p = .684). Performance on incongruent trials during numerical discrimination tasks (i.e., the area salience task) is typically worse relative to congruent trials (DeWind & Brannon, 2012), and this difference in performance is attributed to an attention to surface area in the displays. If learning area symbols heightens the salience of area information as predicted, then participants in the Area Label condition should perform even worse on the incongruent trials of the area salience task at post-test relative to the Area without Labels and Color with Labels Conditions. Moreover, performance should remain consistent from pre- to post-test in the Nonsymbolic Area and Color Label Conditions. In order to determine whether learning domain-specific symbols heightens attention to relevant features in that domain, a 2 (Time: Pre versus Post Test) x 3 (Condition: Area Labels, Nonsymbolic Area, Color Labels) mixed measures ANOVA on the area salience difference score was conducted. There was a main effect of Time, $F(1, 122) = 4.347$, p = .039, $\eta_p^2 =$ = .034, such that there was a greater difference between congruent and incongruent trials at post-test (M = 2.258, SE = .583) compared to pre-test

(M = .900, SE = .532). No other main effects or interactions reached significance, p's > .3, See Figure 5.5.

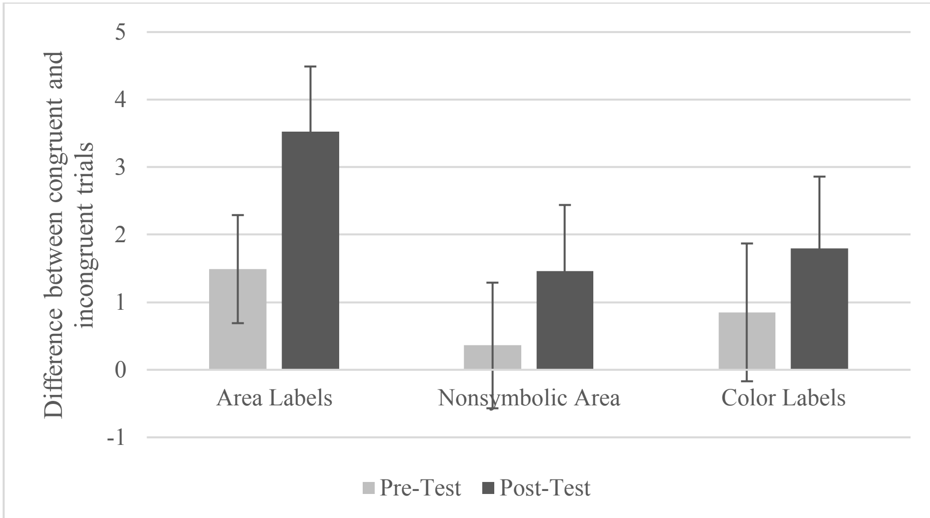

Figure 5.5. Overall, there was a greater difference in incongruent and congruent trials at post-test, such that performance on the incongruent trials worsened over the course of the study.

As an additional test of the effect of learning symbols for area on attention towards area, we next conducted correlations between training performance and another numerical discrimination difference score (difference score at post-test – difference score at pre-test) for each condition separately while also controlling for pre-test performance. None of the correlations reached significance (Area Label: r = -.02, p = .904; Nonsymbolic Area: r = .214, p = .179; Color Label: r = .058, p = .723).

Memory. Memory performance was comparable between the conditions at pre-test, $F(2, 116) = 1.33$, p = .268. We also predicted that learning symbols may enhance memory encoding for surface area. If learning symbols promotes memory-encoding efficiency, then participants who acquired symbols for area should perform better on the memory task at post-test.

To test this, a 2 (Time: Pre versus Post Test) x 3 (Condition: Area Labels, Nonsymbolic Area, Color Labels) mixed measures ANOVA on the accuracy of the Memory task was conducted; however, no main effects or interactions reached significance, *p's* > .1, see Figure 5.6. Although not significant, it is worth noting that performance in the Color Label condition decreased after training, whereas performance in the both Area conditions increased.

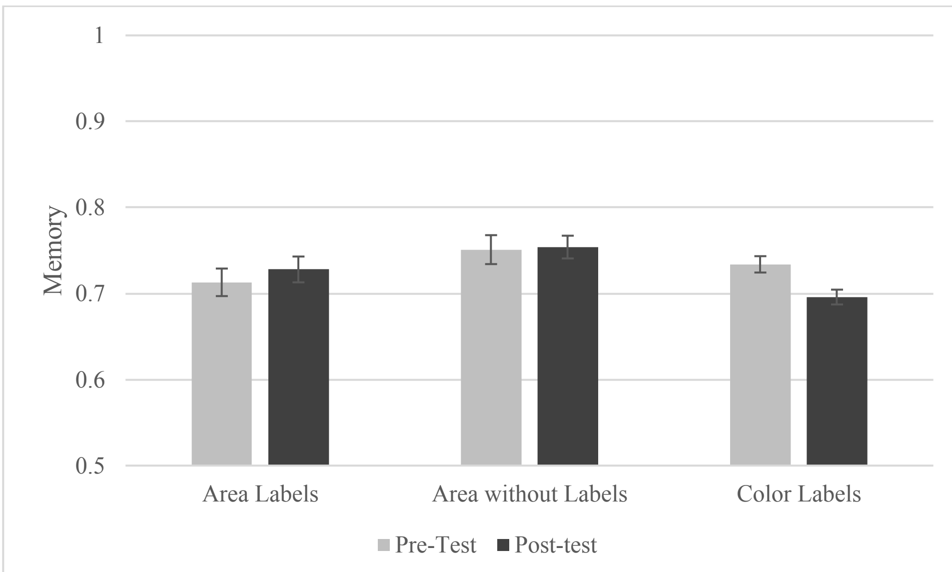

Figure 5.6. Learning symbols for area did not impact memory encoding of spatial dimensions.

We again tested the effect of our training on memory performance by conducting correlations between training performance and a difference score (post-test memory performance – pre-test memory performance), controlling for pre-test performance. Those assigned to the Area Labels condition showed a significant correlation between training and change in memory score, $r = .486$, $p = .002$. However, there were similar correlations were not found in the Nonsymbolic Area ($r = -.061$, $p = .711$) and Color Labels conditions ($r = .080$, $p = .632$). Notably, the difference in correlations was significant between the

Area Labels and Nonsymbolic Area (z = 2.53, p = .005) and between the Area Labels and Color Labels conditions (z = 1.9, p = .03). Correlations were comparable for those assigned to the Nonsymbolic Area and Color Labels conditions (z = -.61, p = .27).

Discussion

The goal of this study was two-fold. Our first aim was to investigate how acquiring symbols for surface area impacts adults' nonsymbolic representations of space. The second objective was to test the mechanism(s) that may underlie enhanced nonsymbolic representations of quantity following symbol acquisition. Based on the language literature (e.g., Boroditsky, 2001), we specifically investigated whether heightened attention and/or enhanced memory encoding might drive this change. Together, this study serves as one of the first experimental investigations of the effect of acquiring symbols for surface area on our nonsymbolic representations of space.

Does learning symbols for surface area enhance nonsymbolic representations of area?

While many studies show relations between an understanding of formal symbols for quantities and the basic processing of those quantities (number: Bonny & Lourenco, 2014; Halberda, Mazzocco, & Feigenson, 2008; Inglis, Attridge, Batchelor, & Gilmore, 2011; Libertus, Feigenson, & Halberda, 2011; Libertus, Odic, & Halberda, 2012; Mazzocco, Feigenson, & Halberda, 2008; Odic, Lisboa, Eisinger, Olivera, Maiche, & Halberda, 2016; Starr, Libertus, & Brannon, 2013; for meta-analysis see: Schneider, Beeres, Coban, Merz, Schmidt, Stricker, de Smedt, 2016; time: Hamamouche & Cordes, accepted), only a single study has experimentally investigated the effect of learning symbols on nonsymbolic representations of quantity (Lindskog, Winaman, & Poom, 2014). Given that adults do not typically learn spatial symbols but instead tend to rely on measurement tools such as rulers when measuring the spatial dimensions of objects, we chose to study this relation in the domain of space so that we could train adults on symbols they were less familiar with (compared to time and number).

Given previous literature showing symbolic knowledge predicted later nonsymbolic acuity (e.g., Lyons et al., 2018; Matejko & Ansari, 2014; Mussolin et al., 2014), we hypothesized that learning symbols for surface area would subsequently enhance performance on a nonsymbolic area discrimination task. Counter to our hypotheses, learning symbols for spatial dimensions did not impact adults' perception of surface area. Although there is evidence of a correlation between symbolic and nonsymbolic spatial acuity (Lauer & Lourenco, 2016; Lourenco & Bonny, 2017), the current data suggest that this relation may not be causal. Why might this be the case? One possibility is that this relation is stronger (and possibly causal) in children who have less experience with spatial symbols. Alternatively, the scaffolding hypothesis may better explain this link, such that nonsymbolic representations form the foundation for symbolic representations, and learning symbols does not necessarily shape nonsymbolic abilities. Additionally, it remains possible that a third variable such as working memory or processing speed may be contributing to this relation.

Rather than improving after symbol instruction, our results showed that overall performance on the area discrimination was *worse* at post-test regardless of condition. Why would learning symbols negatively impact adult's perception of area? One possibility is that participants found the training task quite difficult, and were tired and/or unmotivated to complete the area discrimination task for a second time. This, however, seems unlikely given that overall performance on the training task was quite high and did not differ between the first half and second half of training ($p > .8$). Moreover, performance on our other two post-test measures did not reveal drops in performance relative to pre-test making this explanation doubtful. It remains possible that for some other reason, participants may have been less motivated to complete the area discrimination task a second time, leading to worse performance at post-test.

It is also important to note that participants already performed quite well on the area discrimination task at pre-test leaving little room for improvement. Although we modeled the area discrimination task after previous work (Odic et al., 2013), future research employing more difficult ratios may show that learning symbols for spatial dimensions improves nonsymbolic representations of space. This seems unlikely, however, given that performance on the hardest ratio we tested remained comparable after training ($p > .1$). Additionally, we chose to work with adults who already have some understanding of spatial symbols. While we tried to eliminate previous exposure by using nonsense labels, it is possible that adults' prior knowledge may have played a role in our findings. Children's limited experiences with spatial dimensions, and specifically symbols representing surface area, may allow for greater changes in their nonsymbolic representations of space after symbolic training than that of adults. As such, work with child participants may be a worthwhile pursuit.

Although there was no transfer between the symbolic training and area discrimination performance, we found a relation between performance on the symbolic training and the overall change in area discrimination performance from pre-test to post-test in the Area Labels condition only. This suggests that there may be a link between learning symbols and nonsymbolic representations of those symbols; however, we were unable to tap into this relation in the present study. In future work, one might consider including a more substantial symbolic training that more closely mimics traditional learning of symbols in the classroom either in duration (over several weeks) or using more iterations (more examples of different surface areas) in order to delve more deeply into this relation.

What mechanism(s) might lead to enhanced nonsymbolic representations of area?

We also aimed to explore how learning symbols shapes nonsymbolic representations. Given work from the language literature suggesting that learning symbols may guide attention towards the relevant features (e.g., Boroditsky, 2001; Lupyan & Ward, 2013), enhance memory encoding (e.g., Fausey & Boroditsky, 2011; Schroeder & Marian, 2014), or both, we chose to include an attention (area saliency) and memory measure at pre- and post-test. Although learning symbols for spatial dimensions did not appear to improve nonsymbolic spatial abilities, we still tested whether learning spatial dimensions may have an impact on memory and area salience.

Heightened Attention. Performance on the area salience task indicated that participants may have attended more to surface area of the arrays at post-test, as shown by an overall greater difference in performance on the incongruent and congruent trials at post-test compared to pre-test. However, the effect of Time did not interact with Condition. Thus, counter to predictions, this change over time was not unique to the Area Labels condition, and thus was not a function of learning symbols for spatial dimensions specifically. A distinction in the way in which we assessed participant attention towards spatial dimensions may explain the lack of a Time x Condition interaction. During each training trial, participants attended to a single shape's surface area; during the area salience task, however, participants saw dot arrays. Thus, during the area salience task, participants needed to mentally calculate the cumulative surface area of several dots, or alternatively, determine the average size of an array of heterogeneously-sized dots. In this case, attending to a single item may not necessarily help (or hinder) performance. Thus, we may see a greater effect of training if the assessment of attention also involved a single item like the training, and/or if the dots within the arrays were homogeneous in size. Research using an assessment of attention in which participants needed to attend to the surface area of a single item may result in a more accurate measure of adults' attention to area following the symbolic training.

Memory Encoding. Although we cannot firmly say that attending to area (with or without a label) might enhance memory encoding, we did find a unique relation between training performance and the change in memory performance from pre- to post-test, and this correlation only held in the Area Labels condition. Importantly, the correlation for those assigned to the Area Labels was significantly different than the correlation for those assigned to the Nonsymbolic Area and the Color Labels trainings, suggesting that there may be a unique connection between learning symbols and memory encoding. While these results hint at a possible relation, it is still unclear whether learning area symbols *caused* the change in memory performance. Although we predicted learning area symbols to increase memory encoding for area, it is also possible that individuals with better memory for surface area performed better on the training task. Alternatively, it is possible that a third variable may be at play. For example, greater exposure to surface area through the use of symbols may have supported better memory encoding. Investigating the role that memory plays in symbol acquisition remains an area ripe for investigation.

Additional Explanations. There are many other possible explanations as to why learning symbols did not impact participants' attention to spatial dimensions and/or memory encoding of space. In order to truly test the mechanism(s) impacted during symbol acquisition, we would need to see improvement in adults' area discrimination performance after learning symbols. Given that learning symbols for surface area did not impact adults' perception of area, it is hard to determine what mechanism(s) may have been at play. It is also possible that we were not testing the right mechanism(s) underlying these changes. That is, there may be another means by which learning symbols shapes nonsymbolic representations of quantities. Again, this is difficult to determine given the lack of change in participants' nonsymbolic representations of space. It remains important to study the impact of symbol acquisition on heightened attention and memory encoding.

Conclusions

In conclusion, we found little support for the refinement hypothesis, such that acquiring symbols for surface area did not impact adults' perception of area. Despite this, there was a correlation between performance on the training task and changes in area discrimination performance for those in the Area Labels condition, hinting at a possible relation. Given the current findings, it remains possible that children, who have less experience with symbols, may be more impacted by symbol acquisition. As such, future investigations using a younger age range of participants will be particularly fruitful for further understanding the refinement hypothesis in the domain of space.

Appendix A. Reaction Time

We also used reaction time (RT) as a secondary measure of performance on the area salience and memory tasks. Unfortunately, a programming error prevented us from analyzing RT on the area discrimination task. Scores three standard deviations above or below the mean were excluded from the analyses ($n_{\text{Memory Pre-Test}} = 3$, $n_{\text{Memory Post-test}} = 3$; $N_{\text{Area Salience Difference Score Pre-test}} = 2$, $n_{\text{Area Salience Difference Score Post-test}} = 3$).

Pre-test

Reaction Time revealed significant ratio effects for the area saliency task (Pre-test: $F(3, 387) = 34.003$, $p < .001$, $\eta_p^2 = .209$; Post-test: $F(3, 387) = 81.368$, $p < .001$, $\eta_p^2 = .387$).

Heightened Attention. Reaction time was comparable among the three conditions at pre-test ($F(2, 121) = .409$, $p = .665$). A 2 (Time: Pre-test versus Post-test) x 3 (Condition: Area Labels, Nonsymbolic Area, Color Labels) mixed measures ANOVA was run using reaction time of the numerical discrimination difference score the dependent variable. No main effects or interactions reached significance, *p's* > .5.

Correlational analyses using accuracy on the training and a difference score of the reaction time difference score at post-test minus the difference score at pre-test in each condition separately did not reach significance, (Area Labels, $r = -.163$, $p = .321$; Nonsymbolic Area, $r = -.088$, $p = .595$; Color Labels, $r = -.048$, $p = .767$).

Memory. There were no Condition differences at pre-test ($F(2, 120) = .226$, $p = .798$). We next conducted an identical 2 (Time: Pre-test versus Post-test) x 3 (Condition: Area Labels, Nonsymbolic Area, Color Labels) mixed measures ANOVA using reaction time on the memory task as the dependent variable. There was a main effect of Time, $F(1, 119) = 96.587$, $p < .001$, $\eta_p^2 = .448$, such that participants were quicker at post-test (*M* =

860.32, *SE* =21.59) compared to at pre-test (*M* = 1044.81, *SE* = 29.72). No other main effects or interactions reached significance, *p's* > .2.

Finally, we completed the correlational analyses using accuracy on the training and a difference score on reaction time on the memory task at post-test compared to pre-test. However, none of the correlations reached significance, (Area Labels, $r = -.249$, $p = .116$; Nonsymbolic Area, $r = .169$, $p = .298$; Color Labels, $r = -.295$, $p = .072$).

Appendix B. Covariate Analyses

We also conducted all mixed measures ANOVAs also including pre-test performance as the covariate. This allowed us to test for any effects of pre-test performance on Time and Condition separately. This analysis, however, adjusts average pre-test performance, such that all conditions have identical pre-test performance, making it difficult to interpret Time x Condition interactions, as performance is modified.

When including pre-test area discrimination as a covariate, the a 2 (Time: Pre versus Post Test) x 3 (Condition: Area Labels, Nonsymbolic Area, Color Labels) mixed measures ANCOVA on area discrimination accuracy again revealed a main effect of Time, $F(1, 122) = 12.766$, $p = .001$, $\eta_p^2 = .095$, such that overall performance was worse at post-test ($M = .874$, $SE = .004$) compared to pre-test ($M = .885$, $SE = .000$). There was also a significant Time x Covariate interaction, $F(2, 122) = 13.957$, $p < .001$, $\eta_p^2 = .103$. To further explain the Time x Covariate interaction, we plotted pre-test area discrimination performance on a difference score (post – pre-test performance on the area discrimination task), See Figure A5.1. The negative correlation ($r = -.312$, $p < .001$) indicates that participants who had better performance at pre-test showed less growth than those who had lower performance. No other main effects or interactions reached significance, *p's* > .2.

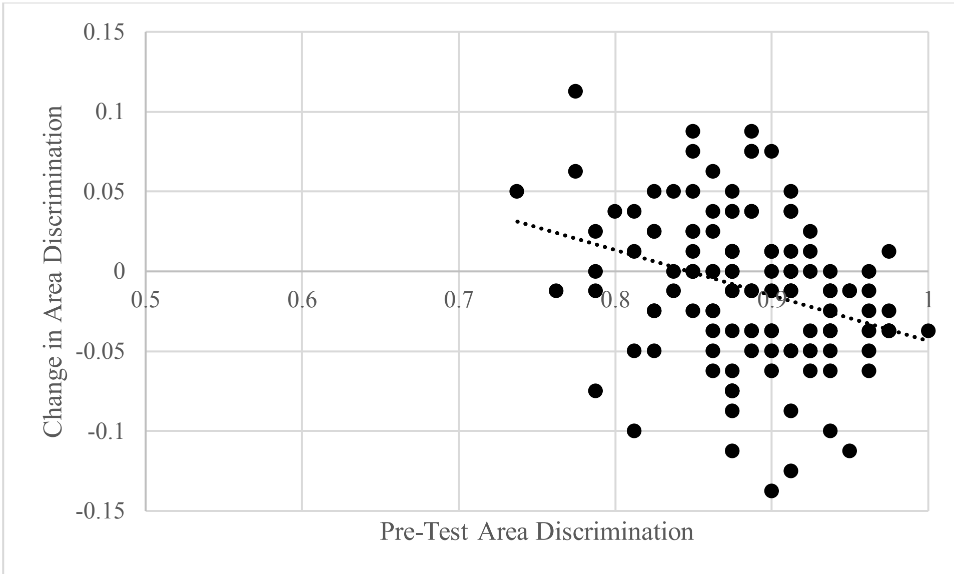

Figure A5.1. Individuals with better area discrimination performance at pre-test showed less change on the area discrimination task over time.

Heightened Attention. A similar pattern of results emerged when including pre-test performance as a covariate in the 2 (Time: Pre versus Post Test) x 3 (Condition: Area Labels, Nonsymbolic Area, Color Labels) mixed measures ANCOVA on the area saliency score. There was again a main effect of Time, $F(1, 121) = 12.009$, $p = .001$, $\eta_p^2 = \; = .090$, such that difference in performance on congruent and incongruent trials was greater at post-test ($M = 2.258$, $SE = .554$) compared to pre-test ($M = .900$, $SE = .000$). There was also a Time x Covariate interaction, $F(1, 121) = 47.519$, $p < .001$, $\eta_p^2 = \; = .282$. The interaction indicated that those who had smaller difference score at pre-test (indicative of comparable performance on the congruent and incongruent trials) showed a larger difference from pre- to post-test ($r = -.524$, $p < .001$), See Figure A5.2. No other main effects or interactions reached significance, *p's* > .3.

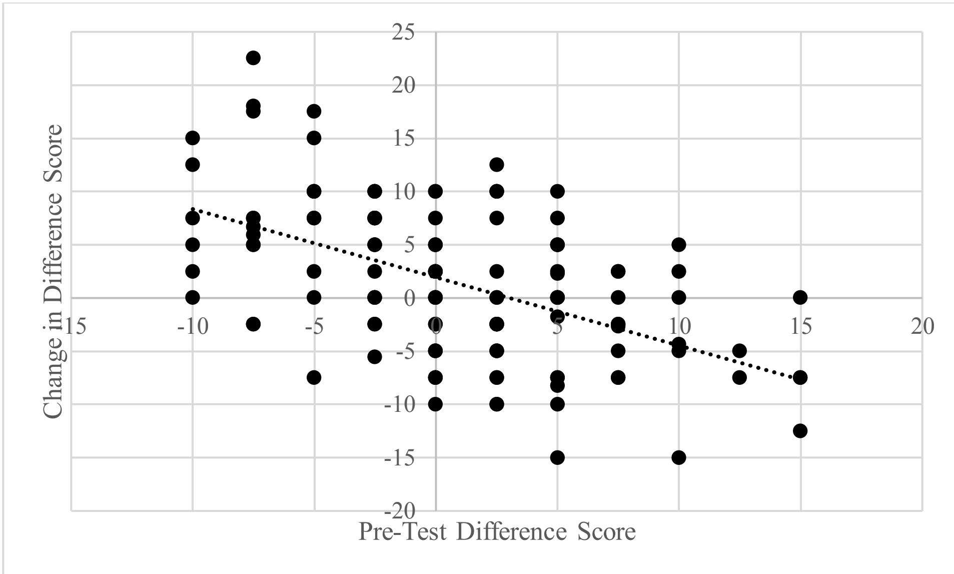

Figure A5.2. Participants with a smaller difference score at pre-test showed more drastic differences between congruent and incongruent trials at post-test.

When including pre-test memory performance as the covariate, 2 (Time: Pre versus Post Test) x 3 (Condition: Area Labels, Nonsymbolic Area, Color Labels) mixed measures ANCOVA on memory performance showed a main effect of Time, $F(1, 114) = 60.658$, $p < .001$, $\eta_p^2 = .347$, such that performance was worse at post-test ($M = .732$, $SE = .007$) compared to pre-test ($M = .735$, $SE = .000$). There was also a main effect of Condition, $F(2, 114) = 3.362$, $p = .038$, $\eta_p^2 = .056$, that was qualified by a significant Time x Condition interaction, $F(2, 114) = 3.358$, $p = .038$, $\eta_p^2 = .056$. While memory performance decreased from pre- to post-test for those assigned to the Color Label condition (Pre-test: $M = .735$, $SE < .001$; Post-test: $M = .705$, $SE = .013$, $p = .023$), those in the Area Labels condition (Pre-test: $M = .735$, $SE <.001$; Post-test: $M = .743$, $SE = .013$, $p = .543$) and Nonsymbolic Area (Pre-test: $M = .735$, $SE < .001$; Post-test: $M = .748$, $SE = .013$, $p = .293$) had comparable performance over time, See Figure A5.3. Finally, there was a Time x Covariate interaction, $F(2, 114) = 62.56$, $p < .001$, $\eta_p^2 = .354$. Those who began with a high memory score at pre-test showed less change at post-test ($r = -.585$, $p < .001$), compared to those participants who began with a lower score, See Figure A5.4.

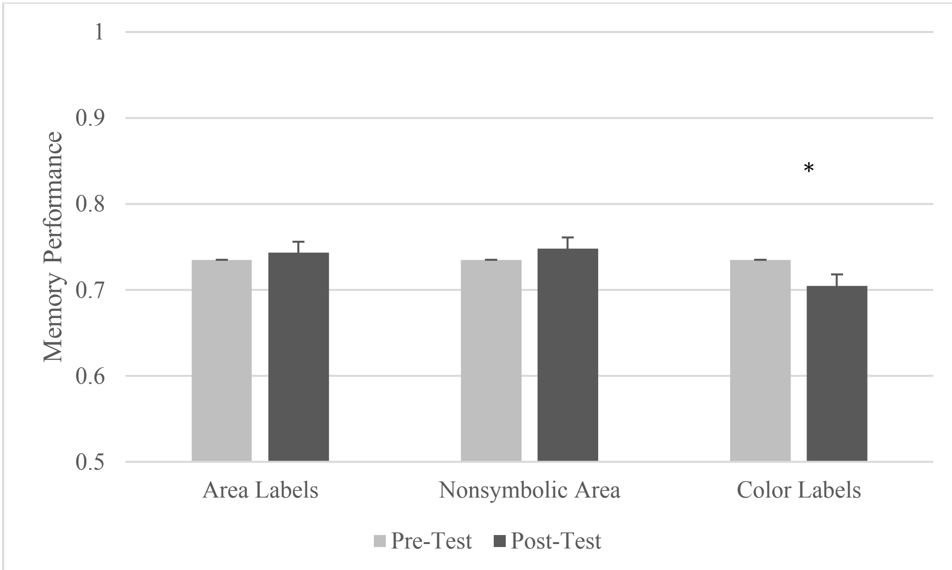

Figure A5.3. Memory performance decreased after training for those in the Color condition.

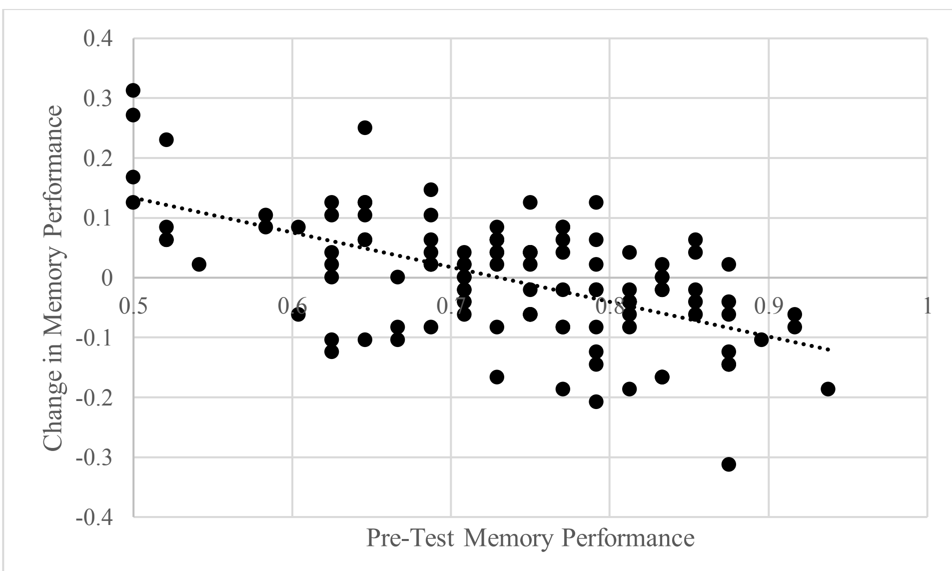

Figure A5.4. Individuals with high memory performance at pre-test showed less change over time compared to individuals with low memory pre-test performance.

CHAPTER 6: IMPLICATIONS AND FUTURE DIRECTIONS

Learning symbols is a critical part of early education. Despite the importance of acquiring symbols to represent quantities, very few studies have investigated how this process impacts our ability to perceive quantities in the absence of language. Does learning symbols for quantities sharpen our nonsymbolic abilities? A growing body of evidence supports this hypothesis by revealing that formal math abilities predict later nonsymbolic numerical acuity (Lyons, Bugden, Zheng, De Jesus, & Ansari, 2018; Matejko & Ansari, 2014; Mussolin, Nys, Content, & Leybaert, 2014; Shusterman, Slusser, Halberda, Odic, 2016; Suárez-Pellicioni & Booth, 2018). While these findings have been taken as evidence in favor of the refinement hypothesis, this work has been largely correlational in nature. As such, experimental manipulations are necessary for truly understanding the causal nature of this relation. To my knowledge, only one experimental investigation has been conducted to investigate the causal nature of this relation. Results revealed no evidence of increased numerical precision after a symbolic number training (Lindskog, Winman, & Poom, 2016). Although this finding discounts the likelihood of the refinement hypothesis, the work was conducted with adults who already had a substantial understanding of numerical symbols. These findings left open the possibility that refinement may occur in the early stages of symbol acquisition, or in children who are still learning symbols through formal education. Moreover, relations between nonsymbolic and symbolic abilities have been limited to the domain of number (e.g., Halberda, Mazzocco, & Feigenson, 2008; for meta-analyses: Chen & Li, 2014; Schneider et al., 2017). Thus, it is unknown whether similar relations hold in the other domains, such as time or space. In other words, does learning temporal units of measurement result in increased temporal acuity? Building upon this literature, the main goal of my dissertation was to employ experimental manipulations to investigate the refinement hypothesis, or the idea that acquiring symbols for quantities shapes our ability to perceive those quantities (e.g., Suárez-Pellicioni & Booth, 2018). I tested this hypothesis in children and adults in the domains of time and space.

Substantial evidence points to a relation between nonsymbolic and symbolic numerical abilities (see Chen & Li, 2014; Schneider et al., 2017) and more recently to a comparable relation in space (Lauer & Lourenco, 2016; Lourenco & Bonny, 2017). Despite this, very few studies have tested whether this relation holds in the domain of time. Given that the refinement hypothesis is contingent upon a relation between symbolic and nonsymbolic abilities, it was critical to first establish this relation in the domain of time. Experiment 1 revealed a relation exists between children's understanding of temporal units of measurement and their temporal acuity as they are learning about temporal symbols through formal education. Importantly, this relation held when controlling for age and numerical abilities, further indicating the uniqueness of the relation. This link was also present in younger children who had yet to receive extensive instruction on temporal symbols in the classroom (Experiment 2). Despite the strength of this relation in childhood, a similar correlation was not present in adults (Experiment 3). Thus, mirroring previous work in the numerical cognition literature, the correlation between nonsymbolic and symbolic timing abilities is only present in childhood.

After establishing this relation, I tested the refinement hypothesis. Surprisingly, I found no evidence in support of the refinement hypothesis in children (Experiment 2) or adults (Experiments 3-4). Specifically, children showed comparable temporal acuity after learning about seconds (Experiment 2). Moreover, modifying adults' symbolic mapping of time did not change their temporal acuity (Experiment 3), nor did learning labels for surface area promote area discrimination performance (Experiment 4). The lack of refinement in adulthood was not particularly shocking given the correlation between adults' symbolic and nonsymbolic timing abilities did not hold; however, it was surprising that children, who have limited symbolic understanding, also showed no improvements. One might suggest that our training was insufficient and did not support symbolic learning. This, however, does not seem to be the case. Although learning symbols did not impact children's temporal acuity, children who learned temporal strategies showed more precise temporal estimates, suggesting that our training was effective. Relatedly, adults immediately adjusted their temporal estimates after receiving feedback, again suggesting that participants acquired some relevant knowledge through training. Thus, although our training improved certain abilities, namely those most closely related to the symbolic manipulation, this improvement did not extend to a subsequent nonsymbolic task. Although the refinement hypothesis was not supported, these studies set the stage for future investigations of this hypothesis, particularly in the domains of time and space.

Implications

Relation between Symbolic and Nonsymbolic Abilities

This dissertation provides the first evidence that a relation exists between nonsymbolic and symbolic temporal abilities. Although a large body of evidence has found this relation in the domain of number (for meta-analyses: Chen & Li, 2014; Schneider et al., 2017), it was unclear whether this relation may be supported by structural isomorphisms between numerical representations. In particular, I predicted that the relation between nonsymbolic and symbolic numerical abilities may be driven by the fact that both representations are discrete in nature. If this were the case, then one would not necessarily expect a similar relation in the domain of time in which discrete temporal symbols must be mapped onto continuous nonsymbolic representations. Despite this, Experiments 1 and 2 revealed a strong relation between nonsymbolic and symbolic temporal abilities. These studies join others showing this relation holds for quantities in which nonsymbolic representations are continuous, such as space (Lauer & Lourenco, 2016; Lourenco & Bonny, 2017). Moreover, this finding suggests that structural isomorphisms are not underlying this relation in the domain of number, and instead indicates that unique relations exist between symbolic and nonsymbolic abilities.

Although we found a strong correlation between children's nonsymbolic and symbolic temporal abilities, this link was not present in adulthood. These findings mirror work in the numerical cognition literature, indicating that this relation tends to be stronger in children than adults (Inglis, Attridge, Batchelor, & Gilmore, 2011). Why would this be the case? Strategy use is the most likely explanation. As we age, adults are more likely to employ strategies that do not necessarily rely on nonsymbolic representations of that quantity. For example, adults tend to use counting strategies when completing a timing task instead of relying on nonsymbolic representations (Fraisse, 1963). Learning these strategies which do not rely on nonsymbolic representations may result in a decreased relation between symbolic and nonsymbolic abilities later in development. Alternatively, adults may develop other domain-specific skills that may mediate this relation. For example, Lyons & Beilock (2011) found symbolic ordering to mediate the relation between adults' nonsymbolic numerical acuity and math performance. Numerical ordering has also been shown to become a better predictor (compared to nonsymbolic acuity) of children's math abilities throughout elementary school (Lyons, Price, Vaessen, Blomert, & Ansari, 2014). While this evidence further supports the idea that nonsymbolic abilities may be a stronger predictor of symbolic abilities early in development, this does not explain previous findings of a similar relation in adulthood (e.g., Agrillo, Piffer, & Adriano, 2013; Libertus, Odic, & Halberda, 2012; Guillaume, Nys, Mussolin, & Content, 2013). Additional research is needed to determine whether this relation exists throughout development, or alternatively, if this relation may only hold prior to developing more advanced strategies.

Refinement Hypothesis

Despite correlational work hinting at the possibility that learning symbols for quantities subsequently refines nonsymbolic acuity (see Lyons et al., 2018; Matejko & Ansari, 2014; Mussolin et al., 2014; Shusterman et al., 2016; Suárez-Pellicioni & Booth, 2018), the experimental manipulations we employed revealed no evidence for this hypothesis. Although surprising, there are many possible explanations for this finding. Firstly, it is possible that counter to the refinement hypothesis, learning symbols does not shape nonsymbolic abilities. In this instance, the cause of improving nonsymbolic abilities over development remains an open question (see Future Directions). Alternatively, it is possible that our training was not sufficient enough to impact nonsymbolic abilities. There is evidence to suggest that our trainings increased symbolic knowledge. For example, children showed more precise temporal estimates following symbolic time training, and adults shifted their symbolic mapping of time after feedback. These adjustments, however, did not extend to a subsequent nonsymbolic task. Thus, it remains unknown whether the trainings were sufficient to impact subsequent nonsymbolic abilities. A longer and more intense symbolic training may further augment symbolic knowledge and consequently result in changes in nonsymbolic abilities. Alternatively, it is possible that this causal relation may be specific to the domain of number. While possible, it seems unlikely given that the only experimental test of the refinement hypothesis reported no evidence of improved nonsymbolic numerical acuity after a symbolic math training (Lindskog et al., 2016). While the current dissertation serves as one of the first systematic investigations of the refinement hypothesis, substantially more work is needed to fully understand whether a causal relation between symbolic and nonsymbolic abilities exists, and whether this relation is comparable across different quantity domains.

In addition to testing the refinement hypothesis, I also intended to investigate the impact of symbol acquisition on cognitive mechanisms that may have supported improved nonsymbolic acuity. Specifically, I predicted that learning symbols may lead to enhanced memory encoding and/or heightened attention to quantity dimensions (see Boroditsky, 2001). To test this, I measured adults' attention to and memory of spatial dimensions before and after learning symbols to represent surface area. Results revealed no difference in attention towards spatial cues after symbol acquisition. However, the data hinted at the possibility that learning symbols may enhance memory encoding, although the effect was not particularly strong. It remains possible that a more substantial training, in which refinement occurs, may point to the mechanism(s) underlying these changes. Moreover, it may be fruitful to investigate the cognitive mechanism(s) underlying changes in nonsymbolic abilities in children who are still in the process of acquiring units of measurement. Finally, it is plausible that mechanism(s) other than the ones chosen in the present study are implicated during symbol acquisition.

Future Directions

What causes improvement in nonsymbolic abilities?

It is well documented that nonsymbolic abilities improve over development (Droit-Volet, Clément, & Fayol, 2008; Odic, 2018; for a review: Feigenson, 2007). However, the cause of improvement has been left to speculation. In this dissertation, I tested the possibility that acquiring symbols for quantities results in enhanced nonsymbolic abilities. Counter to my predictions, learning symbols did not appear to impact nonsymbolic abilities. What causes improvements in nonsymbolic acuity if not acquiring symbols? Related to the refinement hypothesis, development of general linguistic abilities may mold nonsymbolic abilities (see Posid & Cordes, 2015; Shusterman et al., 2016). However, this account does not explain increases in temporal and numerical precision in infancy (for a review: Feigenson, 2007), as this is a time at which language has yet to be acquired. General maturation has also been suggested as the mechanism underlying these changes, although this prediction is harder to test as general maturation is confounded with many other developments. Given evidence that nonsymbolic abilities continue developing into adulthood (Odic, 2018), general maturation remains a valid explanation. It is also possible that general maturation is responsible for nonsymbolic development in infancy; however, burgeoning language abilities and/or general experience with quantities may support improvements later in development.

A slightly less interesting and less explored possibility is that improving domain general abilities may cause changes in nonsymbolic acuity. There is substantial work suggesting that working memory abilities are related to temporal acuity (e.g., Broadway & Engle, 2011; Odic et al., 2016; Zélanti & Droit-Volet, 2011). For example, Zélanti and Droit-Volet (2011) found that age related changes in sub-second timing were completely explained by working memory abilities. Processing speed may also play a role in nonsymbolic acuity. Work in favor of this possibility demonstrates that both processing speed and working memory are important predictors of temporal acuity (Droit-Volet & Zélanti, 2013). This work, however, has been limited to correlational studies. Although relations exist between domain-general abilities and nonsymbolic acuity, it remains unknown whether changes in domain-general abilities *cause* changes in nonsymbolic abilities. Future work employing experimental manipulations will be crucial for determining whether domain general abilities may account for enhanced nonsymbolic processing throughout development. In particular, it may be beneficial to see whether manipulating domain-general abilities subsequently impacts nonsymbolic acuity.

Further Investigations of the Refinement Hypothesis

While this dissertation serves as one of the first experimental manipulations of the refinement hypothesis, the topic remains rich for future investigation. Although we tested children and adults' abilities before and after learning temporal or spatial symbols, our symbolic training was relatively brief. In particular, children's symbolic training lasted no more than 20 minutes, and adults' training was even shorter. Throughout formal education, learning symbols is an ongoing activity that lasts several months, or even years. That is, although children are born with the ability to represent quantities nonsymbolically, acquiring the cardinal principle (a proxy for symbolic numerical abilities) and learning the meaning of temporal words (e.g., "tomorrow", "next week", "yesterday") takes children several years to master (e.g., time: Tillman & Barner, 2015; Tillman, Marghetis, Barner, & Srinivasan, 2017; number: Wynn, 1992). It is possible that trainings which more closely mimic typical symbol instruction may be more likely to elicit changes to nonsymbolic abilities. That is, we may be more likely to find support for the refinement hypothesis after teaching participants about symbols in a way that more closely reflects traditional learning. Future research employing more extensive symbol instruction may be critical for further testing the refinement hypothesis. Specifically, it may be beneficial to track children's nonsymbolic acuity over the course of several weeks of symbolic training in a more formal setting.

In addition to lengthening the training, future work may also investigate how acquiring different types of symbolic knowledge may relate to nonsymbolic abilities. In our time training, for example, we specifically taught children about seconds, as opposed to minutes or hours. Similarly, in the experiment with adults, we targeted timing in the millisecond range. I predicted that learning the duration of a second/millisecond may best inform symbolic timing abilities by lessening the noise in symbolic representations. That is, prior to training children may have estimated anything between 1-3 seconds as being 2 seconds. After learning the specific duration of 2 seconds, however, children may have a more precise understanding of that duration (i.e., lasting 1.5-2.5 seconds). Moreover, because the temporal discrimination task took place in the milliseconds and seconds range, I reasoned that learning about seconds/milliseconds would be most likely to refine temporal acuity. These concepts (i.e., milliseconds, seconds, etc.) however, are rarely, if ever, formally taught in the classroom. Instead, the focus tends to be on learning the relations between minutes and hours and telling time on clocks (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). Being able to read clocks, however, does not necessarily reflect symbolic representations of time, nor would acquiring this skill inevitably lessen the noise in symbolic representations. While it seems unlikely that learning how to tell time and/or read clocks would enhance one's temporal acuity, it would be interesting to determine whether typical classroom instruction improves nonsymbolic abilities. Testing this possibility could take many forms. For example, tracking children's temporal acuity throughout 1st and 2nd grade when they are learning temporal units of measurement in the classroom would be one method of investigation. Alternatively, in collaboration with educators, researchers could design a training that teachers would implement in the classroom. The training would include timing concepts traditionally taught in the classroom, in addition to, abilities tapping directly into symbolic representations of time. Children's nonsymbolic abilities would be tested before and after implementation in order to determine whether learning about temporal symbols in a more traditional learning setting might promote temporal acuity.

Conclusions

In conclusion, I contend that the relation between symbolic and nonsymbolic representations of quantity may be stronger in children who are still in the process of acquiring symbolic knowledge compared to adults. Moreover, although acquiring symbols did not impact subsequent nonsymbolic acuity, I propose many future directions to further test the likelihood of the refinement hypothesis. The present data also hint at the possibility that memory encoding, but not heightened attention may underlie changes following symbol acquisition. Given the current findings, several open questions remain: If not learning symbols, what causes improvements in nonsymbolic acuity over development? How much symbolic knowledge is needed before refinement occurs, if ever? Does the refinement hypothesis hold at specific points in development? Together these data shed light on the refinement hypothesis in time and space, while also pointing to exciting areas for future investigation.

REFERENCES

Agrillo, C., Piffer, L., & Adriano, A. (2013). Individual differences in non-symbolic numerical abilities predict mathematical achievements but contradict ATOM. *Behavioral and Brain Functions*, *9*(1), 26.

Ames, L. B. (1946). The development of the sense of time in the young child. *The Pedagogical Seminary and Journal of Genetic Psychology*, *68*(1), 97-125.

Athanasopoulos, P., Wiggett, A., Dering, B., Kuipers, J. R., & Thierry, G. (2009). The Whorfian mind: Electrophysiological evidence that language shapes perception. *Communicative & integrative biology*, *2*(4), 332-334.

Arlin, M. (1990). What happens to time when you sleep? Children's development of objective time and its relation to time perception. *Cognitive Development*, *5*(1), 71-88.

Babai, R., Nattiv, L., & Stavy, R. (2016). Comparison of perimeters: improving students' performance by increasing the salience of the relevant variable. *ZDM, 48(3),* 367-378.

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants?. *Journal of experimental child psychology*, *64*(1), 3-26.

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(39), 14116-14121.

Bonny, J. W., & Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: Evidence from the preschool years. *Journal of experimental child psychology*, *114*(3), 375-388.

Bonny, J. W., & Lourenco, S. F. (2015). Individual differences in children's approximations of area correlate with competence in basic geometry. *Learning and Individual Differences*, *44*, 16-24.

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive psychology*, *43*(1), 1-22.

Boroditsky, L. (2009). How does our language shape the way we think. *What's next*, 116-129.

Boroditsky, L. (2011). How language shapes thought. *Scientific American*, *304*(2), 62-65.

Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently?. *Cognition*, *118*(1), 123-129.

Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision: A label advantage in object recognition. *Journal of Neuroscience*, *35*(25), 9329-9335.

Boyer, T. W., & Levine, S. C. (2015). Prompting children to reason proportionally: Processing discrete units as continuous amounts. *Developmental psychology, 51(5),* 615.

Bradley, N. C. (1947). THE GROWTH OF THE KNOWLEDGE OF TIME IN CHILDREN OF SCHOOL-AGE. *British Journal of Psychology*, *38*(2), 67-78.

Broadway, J. M., & Engle, R. W. (2011). Individual differences in working memory capacity and temporal discrimination. *PloS One, 6*(10), e25422.

Brannon, E. M., Lutz, D., & Cordes, S. (2006). The development of area discrimination and its implications for number representation in infancy. *Developmental science*, *9*(6), F59-F64.

Brannon, E. M., Suanda, S., & Libertus, K. (2007). Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. *Developmental Science, 10*(6), 770-777.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282*(5389), 746-749.

Bueti, D., Walsh, V., Frith, C., & Rees, G. (2008). Different brain circuits underlie motor and perceptual representations of temporal intervals. *Journal of Cognitive Neuroscience*, *20*(2), 204-214.

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*(2), 83-91.

Casasanto, D. (2008). Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Language learning, 55*(s1), 63-79.

Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S., ... & Gil, D. (2004, January). How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish. In *Proceedings of the Cognitive Science Society* (Vol. 26, No. 26).

Casini, L., Pech-Georgel, C., & Ziegler, J. C. (2017). It's about time: revisiting temporal processing deficits in dyslexia. *Developmental science*.

Chen, Q. & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta psychologica, 148,* 163-172.

Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M., & Burr, D. C. (2012). Optimal encoding of interval timing in expert percussionists. *Journal of Neuroscience*, *32*(3), 1056-1060.

Clayton, S., & Gilmore, C. (2015). Inhibition in dot comparison tasks. *Zdm*, *47*(5), 759-770.

Cohen-Mimran, R., & Sapir, S. (2007). Auditory temporal processing deficits in children with reading disabilities. *Dyslexia*, *13*(3), 175-192.

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. Cognition, 114(3), 356-371.

Danckert, J. A., & Allman, A. A. A. (2005). Time flies when you're having fun: Temporal estimation and the experience of boredom. *Brain and cognition*, *59*(3), 236-245.

DeWind, N. K., & Brannon, E. M. (2012). Malleability of the approximate number system: effects of feedback and training. *Frontiers in human neuroscience*, *6*, 68.

Droit-Volet, S. (2013). Time perception in children: A neurodevelopmental approach. *Neuropsychologia*, *51*(2), 220-234.

Droit-Volet, S., Clément, A., & Fayol, M. (2008). Time, number and length: Similarities and differences in discrimination in adults and children. *Quarterly Journal of Experimental Psychology*, *61*(12), 1827-1846.

Droit-Volet, S., & Wearden, J. H. (2001). Temporal bisection in children. *Journal of Experimental Child Psychology*, *80*(2), 142-159.

Droit-Volet, S., & Zélanti, P. S. (2013a). Development of time sensitivity and information processing speed. *PloS one*, *8*(8), e71424.

Droit-Volet, S., & Zélanti, P. (2013b). Development of time sensitivity: duration ratios in time bisection. *The Quarterly Journal of Experimental Psychology*, *66*(4), 671-686.

Emerson, R. W., & Cantlon, J. F. (2015). Continuity and change in children's longitudinal neural responses to numbers. *Developmental science*, *18*(2), 314-326.

Espy, K. A. (1997). The Shape School: Assessing executive function in preschool children. *Developmental Neuropsychology*, *13*(4), 495-499.

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. Psychonomic bulletin & review, 2(4), 460-493.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160.

Fausey, C. M., & Boroditsky, L. (2011). Who dunnit? Cross-linguistic differences in eye-witness memory. *Psychonomic bulletin & review*, *18*(1), 150-157.

Feigenson, L. (2007). The equality of quantity. *Trends in cognitive sciences*, *11*(5), 185-187.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in cognitive sciences*, *8*(7), 307-314.

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Development Perspectives*, *7*(2), 74-79.

Fraise, P. (1963). The psychology of time.

Friedman, W. J. (1977). The development of children's understanding of cyclic aspects of time. *Child Development*, 1593-1599.

Friedman, W. J., & Laycock, F. (1989). Children's analog and digital clock knowledge. *Child Development*, 357-371.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge: The MIT Press.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*(3), 279.

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., ... & Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PloS one*, *8*(6), e67374.

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition*, *115*(3), 394-406.

Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception & psychophysics, 54*(3), 383-394.

Grondin, S., Meilleur-Wells, G., & Lachance, R. (1999). When to start explicit counting in a time-intervals discrimination task: A critical point in the timing process of humans. *Journal of experimental psychology: human perception and performance, 25*(4), 993.

Guillaume, M., Nys, J., Mussolin, C., & Content, A. (2013). Differences in the acuity of the Approximate Number System in adults: The effect of mathematical ability. *Acta psychologica*, *144*(3), 506-512.

Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665-668.

Hamamouche, K. & Cordes, S. (2019). A Divergence of Sub- and Supra-Second Timing in Childhood and its Relation to Academic Achievement. *Journal of experimental* *child psychology, 178,* 137-154.

Hamamouche, K., Keefe, M., Jordan, K., & Cordes, S. (2018). Cognitive Load Affects Numerical and Temporal Judgments in Distinct Ways. *Frontiers in psychology, 9*.

Harrison, M. L. (1934). The Nature and Development of Concepts of Time among Young Children. *The Elementary School Journal, 34*(7), 507-514.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of experimental child psychology*, *103*(1), 17-29.

Hood, M., & Conlon, E. (2004). Visual and auditory temporal processing and early reading development. *Dyslexia*, *10*(3), 234-252.

Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences*, *103*(51), 19599-19604.

Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition, 131(*1), 92-107.

Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. *Psychonomic Bulletin & Review*, *18*(6), 1222-1229.

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382-10385.

Kramer, P., Bressan, P., & Grassi, M. (2011). Time estimation predicts mathematical intelligence. *PloS one*, *6*(12), e28621.

Lauer, J. E., & Lourenco, S. F. (2016). Spatial processing in infancy predicts both spatial and mathematical aptitude in childhood. *Psychological science*, *27*(10), 1291-1298.

LeFevre, J. A., Skwarchuk, S. L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *41*(2), 55.

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, *14*(6), 1292-1300.

Libertus, M. E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta psychologica*, *141*(3), 373-379.

Lindskog, M., Winman, A., & Poom, L. (2016). Arithmetic training does not improve approximate number system acuity. *Frontiers in psychology*, *7*, 1634.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science*, *14*(5), 396-401.

Lourenco, S. F., & Bonny, J. W. (2017). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Developmental Science*, *20*(4), e12418.

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences, 109*(46), 18737-18742.

Lupyan, G. (2008). The conceptual grouping effect: Categories matter (and named categories matter more). *Cognition*, *108*(2), 566-577.

Lupyan, G., & Spivey, M. J. (2010). Making the invisible visible: Verbal but not visual cues enhance visual detection. *PloS one*, *5*(7), e11452.

Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences*, *110*(35), 14196-14201.

Lusardi, A. (2012). Numeracy, financial literacy, and financial decision-making (No. w17821). National Bureau of Economic Research.

Lustig, C., & Meck, W. H. (2011). Modality differences in timing and temporal memory throughout the lifespan. *Brain and cognition*, *77*(2), 298-303.

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition*, *121*(2), 256-261.

Lyons, I. M., Bugden, S., Zheng, S., De Jesus, S., & Ansari, D. (2018). Symbolic number skills predict growth in nonsymbolic number skills in kindergarteners. *Developmental psychology, 54*(3), 440.

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental science*, *17*(5), 714-726.

Matejko, A. A., & Ansari, D. (2016). Trajectories of symbolic and nonsymbolic magnitude processing in the first year of formal schooling. *PloS one*, *11*(3), e0149863.

Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011a). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS* *one*, *6*(9), e23749.

Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011b). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development*, *82*(4), 1224-1237.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(3), 320.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, *215*(5109), 1519-1520.

Mussolin, C., Nys, J., Content, A., & Leybaert, J. (2014). Symbolic number abilities predict later approximate number system acuity in preschool children. *PLoS One*, *9*(3), e91839.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). Common Core State Standards Mathematics. National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington, D.C.

Nys, J., & Content, A. (2012). Judgement of discrete and continuous quantity in adults: Number counts!. *The Quarterly Journal of Experimental Psychology*, *65*(4), 675-690.

Oakden, E. C. & Sturt, M. (1922). The Development of the Knowledge of Time in Children. *British Journal of Psychology.*

Odic, D. (2018). Children's intuitive sense of number develops independently of their perception of area, density, length, and time. *Developmental Science*.

Odic, D., Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Developmental Psychology*, *49*(6), 1103.

Odic, D., Lisboa, J. V., Eisinger, R., Olivera, M. G., Maiche, A., & Halberda, J. (2016). Approximate number and approximate time discrimination each correlate with school math abilities in young children. *Acta Psychologica, 163,* 17-26.

Ostarek, M., & Huettig, F. (2017). Spoken words can make the invisible visible—Testing the involvement of low-level visual representations in spoken word processing. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(3), 499.

Park, J., & Brannon, E. M. (2016). How to interpret cognitive training studies: A reply to Lindskog & Winman. *Cognition*, *150*, 247-251.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological science, 17*(5), 407-413.

Piazza, M., Pica, P., Izard, V., Spelke, E. S., & Dehaene, S. (2013). Education enhances the acuity of the nonverbal approximate number system. *Psychological Science*, *24*(6), 1037-1043.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499-503.

Platt, J. R., & Davis, E. R. (1983). Bisection of temporal intervals by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 9(2),* 160.

Plourde, M., Gamache, P. L., Laflamme, V., & Grondin, S. (2017). Using time-processing skills to predict reading abilities in elementary school children. *Timing & Time*

Perception, *5*(1), 35-60.

Posid, T., & Cordes, S. (2015). Verbal counting moderates perceptual biases found in children's cardinality judgments. *Journal of Cognition and Development*, *16*(4), 621-637.

Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta psychologica*, *140*(1), 50-57.

Provasi, J., Rattat, A. C., & Droit-Volet, S. (2011). Temporal bisection in 4-month-old infants. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(1), 108.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological bulletin, 135*(6), 943.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science,* 1(6), 906-914.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 1926-1928.

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., De Smdedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental science, 20(3),* e12372.

Schroeder, S. R., & Marian, V. (2014). Bilingual episodic memory: How speaking two languages influences remembering. In *Foundations of bilingual memory* (pp. 111-132). Springer, New York, NY.

Shatz, M., Tare, M., Nguyen, S., & Young, T. (2010). Acquiring Non-Object Terms: The Case for Time Words. *Cognitive Development 11*(1), 16-36.

Shusterman, A., Slusser, E., Halberda, J., & Odic, D. (2016). Acquisition of the cardinal principle coincides with improvement in approximate number system acuity in preschoolers. *PloS one*, *11*(4), e0153072.

Skagerlund, K., & Träff, U. (2016). Processing of space, time, and number contributes to mathematical abilities above and beyond domain-general cognitive abilities. *Journal of Experimental Child Psychology*, *143*, 85-101.

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, *110*(45), 18116-18120.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, *64*(3), 153.

Suárez-Pellicioni, M., & Booth, J. R. (2018). Fluency in symbolic arithmetic refines the approximate number system in parietal cortex. *Human brain mapping*, *39*(10), 3956-3971.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*(2), 182-198.

Thomas, E. C., & Brown, I. (1974). Time perception and the filled-duration illusion. *Perception & Psychophysics*, *16*(3), 449-458.

Tillman, K. A., & Barner, D. (2015). Learning the language of time: Children's acquisition of duration words. *Cognitive Psychology*, *78*, 57-77.

Tillman, K. A., Marghetis, T., Barner, D., & Srinivasan, M. (2017). Today is tomorrow's yesterday: Children's acquisition of deictic time words. *Cognitive psychology, 92,* 87-100.

Tosun, S., Vaid, J., & Geraci, L. (2013). Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language, 69*(2), 121-134.

VanMarle, K., & Wynn, K. (2006). Six-month-old infants use analog magnitudes to represent duration. *Developmental science*, *9*(5), F41-F49.

van Marle, K., & Wynn, K. Quantitative Reasoning. *Encyclopedia of Cognitive Science*.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483-488.

Wang, J. J., Odic, D., Halberda, J., & Feigenson, F. (2016). Changing the precision of preschoolers' approximate number system representations changes their symbolic math performance. *Journal of Experimental Chil Psychology, 147* 82-99.

Wearden, J. H., Norton, R., Martin, S., & Montford-Bebb, O. (2007). Internal clock processes and the filled-duration illusion. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(3), 716.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, *104*(19), 7780-7785.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive psychology*, *24*(2), 220-251.

Xu, F., & Arriaga, R. I. (2007). Number discrimination in 10-month-old infants. *British* *journal of developmental psychology, 25*(1), 103-108.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1-B11.

Zélanti, P. S., & Droit-Volet, S. (2011). Cognitive abilities explaining age-related changes in time perception of short and long durations. *Journal of Experimental Child Psychology*, *109*(2), 143-157.